LACUS FORUM XXVII

Speaking and Comprehending

RICE

© 2009 The Linguistic Association of Canada and the United States (LACUS). The content of this volume is from *LACUS Forum* 27 (published 2001).
This volume of the LACUS Forum is being made available by the Linguistic Association of Canada and the United States under the Creative Commons Attribution-NonCommercial 3.0 license. See below.

YOUR RIGHTS
This electronic copy is provided free of charge with no implied warranty. It is made available to you under the terms of the Creative Commons Attribution-NonCommercial license version 3.0 (*http://creativecommons.org/licenses/by-nc/3.0/*)

Under this license you are free:
- **to Share** — to copy, distribute and transmit the work
- **to Remix** — to adapt the work

Under the following conditions:
- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.

With the understanding that:
- **Waiver** — Any of the above conditions can be waived if you get permission from the copyright holder.
- **Other Rights** — In no way are any of the following rights affected by the license:
  - Your fair dealing or fair use rights;
  - The author's moral rights;
  - Rights other persons may have either in the work itself or in how the work is used, such as publicity or privacy rights.

Notice: For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to the web page cited above.
For inquiries concerning commercial use of this work, please visit *http://lacus.weebly.com/publications.html*

Cover: The front cover of this document is licensed under the Creative Commons Attribution-No Derivative Works 3.0 license (*http://creativecommons.org/licenses/by-nd/3.0/*) and may not be altered in any fashion. The LACUS "lakes" logo and Rice University logo on the cover are trademarks of LACUS and Rice University, respectively. The Rice University logo is used here with permission from the trademark holder. No license for use of these trademarks outside of redistribution of this exact file is granted. These trademarks may not be included in any adaptation of this work.

Contents

I. Presidential and Prize-Winning Lectures

Presidential Address: An Integrated Neurolinguistic Theory of Bilingualism 3
Michel Paradis
Presidents' Post-Doctoral Prize: Particle Placement in English 17
Stefan Th. Gries
Presidents' Pre-Doctoral Prize: Nobility in the History of Adjectives Ending in *-ic* and *-ical* 33
Mark Kaunisto

II. Lexical Studies

French *Ça* and the Dynamics of Reference 49
Michel Achard
What's in the Head? Metaphorical Expressions in Polish and English 63
Barbara Bacz
Data and Science in Introductory Linguistics Textbooks 75
Douglas W. Coleman
Verbs of Tolerance with the Gerund and the Infinitive 87
Patrick J. Duffley
Fashionable *You* 93
Connie Eble
When Do Sciences Make Nouns from Verbs, and When Don't They? 103
Carolyn G. Hartnett
Lexicalization and Opacity 113
Aya Katz
An Operator-Based Integration of Comprehension and Generation 123
Deryle Lonsdale
Given vs. New Information Influencing Constituent Ordering in the Verb-Particle Construction 133
Julia Peters
Does Synchrony Recapitulate Diachrony? A Case Study From Japanese 141
Gary D. Prideaux
Categorizing English Superposition 151
Bernard Paul Sypniewski

III. Discourse Studies

Iteration and Genre in a 19th-Century Novel 163
Michael Cummings
Verbal Reasoning and Brain Hemispheres 175
Inga B. Dolinina
Presidents' Pre-Doctoral Commendation: Discourse Markers—Accessing 'Hearer-Old' Information: The Case of Russian *Že* 187
Anna Feldman
Ancient Theorizing on Nonverbal Communication 203
Thorsten Fögen
The Grammar of the Pictish Symbol Stones 217
Toby D. Griffen
Presidents' Pre-Doctoral Commendation: Episode Transitions and their Linguistic Encodings 227
Shaojun Ji
Poetic License, Embellished by Music 237
Saul Levin
Explicit/Implicit Modularity in Right Hemisphere Pragmatics 247
Jennifer Mortimer
The Amplification Clause Root in San Miguel El Grande Mixtec 259
Evelyn G. Pike

IV. Phonological Studies

Does Perception Precede Production? Evidence from Korean-English Bilinguals 273
Wendy Baker & Pavel Trofimovich
Vowel Discrimination by Speakers of English, German, Japanese, Mandarin and Cantonese 285
David C. Bennett, Zhang Jia-Lu, Lü Shi-Nan, Hu Xing-Hui, Andris Abakuks & Valérie Hazan
Perception and Production in a Series of Related Mergers 297
David Bowie
Phoneme and Grapheme: How Parallel Can They Be? 307
David G. Lockwood
Deriving OT Constraints From Relational Network Theory 317
William J. Sullivan
Syllabification in Parallel in Spanish 327
Caroline R. Wiltshire

2000 PRESIDENTIAL ADDRESS

AN INTEGRATED NEUROLINGUISTIC THEORY OF BILINGUALISM (1976–2000)

Michel Paradis
McGill University

INTRODUCTION. My very first paper published on bilingualism appeared in the *Third LACUS Forum* (Paradis 1977). It was called 'The stratification of bilingualism'. It contained the germ of the Three-Store Hypothesis and of the Subset Hypothesis, though these did not clearly occur to me as such until the *Sixth* and *Seventh LACUS Forum*, respectively. The main point of that first paper, presented at the meeting in El Paso in 1976, was that each level of grammar of each of a bilingual's languages could independently contain various degrees of unidirectional or bidirectional interference, in keeping with Weinreich's (1953) coordinate, compound and subordinate organization. This paper also contained the seed of the idea, made more explicit later, that the *contents* of the representation (i.e., whether it contains illegitimate features of another language or not) are orthogonal to the representation of two languages in one brain. It dealt with the *linguistic* level of description—a description of the grammar of a given bilingual individual at a given time (Figure 1, overleaf). There were two main points: (1) the language system of a bilingual individual may contain, in various proportions at each stratal level (roughly corresponding to phonology, morphology, syntax and lexical semantics), all three types of interference (i.e., unidirectional, bidirectional, or none); (2) the state of bilingualism is not static, as the various interlingual relationships will change, over time, possibly differently at each stratum, as a result of experience.

The Three-Store Hypothesis, as it was initially proposed in 1979 at the sixth LACUS Forum in Calgary, was a *psycholinguistic* construct, separating the nonlinguistic cognitive store from the two language stores of bilingual speakers. It was meant to reconcile the contradictory findings of experimental psycholinguists of the 1960s and 1970s who found evidence, now in support of one store, now of two stores (Figure 2a, overleaf). Kolers (1968) reported that his subjects' responses were neither fully compatible with the one-store nor with the two-store account, but rather were similar to responses to synonyms in the same language. This led to the proposal of the Three-Store Hypothesis (Figure 2b, overleaf). The Subsystems Hypothesis, first presented here at Rice University in 1980, and published in the *Seventh LACUS Forum* (Paradis 1981), was a *neurofunctional* proposal, compatible with all the various recovery patterns of bilingual aphasic patients, including alternating antagonism and selective aphasia, which were described for the first time some years later (Paradis et al. 1982; Paradis & Goldblum 1989; Nilipour & Ashayeri 1989). It too assumes a separate neurofunctional cognitive system independent of the language system, and hence of the language systems of bilinguals. In addition, it assumes that language is an independent neurofunctional system, a neurofunctional module, receiving inputs

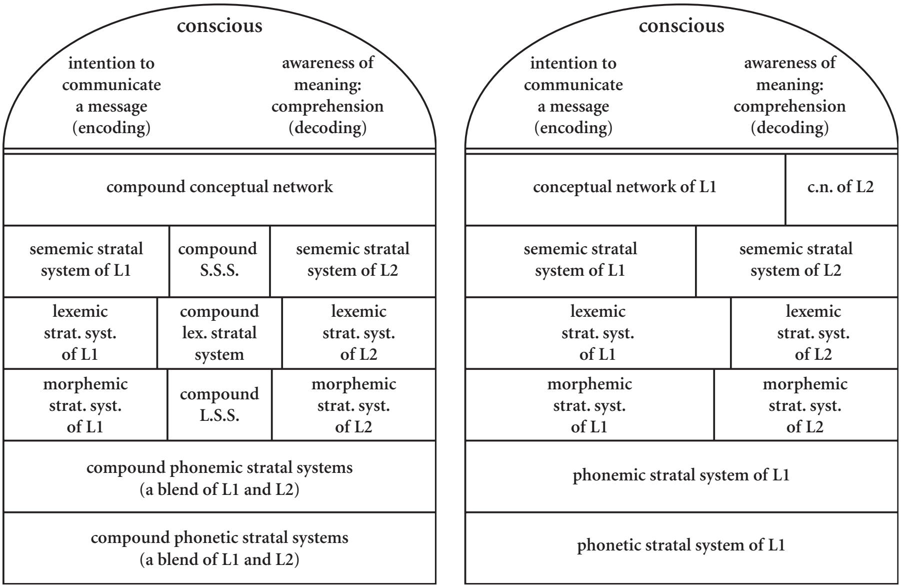

conscious
intention to communicate a message (encoding)
awareness of meaning: comprehension (decoding)
compound conceptual network
sememic stratal system of L1
compound S.S.S.
sememic stratal system of L2
lexemic strat. syst. of L1
compound lex. stratal system
lexemic strat. syst. of L2
morphemic strat. syst. of L1
compound L.S.S.
morphemic strat. syst. of L2
compound phonemic stratal systems (a blend of L1 and L2)
compound phonetic stratal systems (a blend of L1 and L2)
conscious
intention to communicate a message (encoding)
awareness of meaning: comprehension (decoding)
conceptual network of L1
c.n. of L2
sememic stratal system of L1
sememic stratal system of L2
lexemic strat. syst. of L1
lexemic strat. syst. of L2
morphemic strat. syst. of L1
morphemic strat. syst. of L2
phonemic stratal system of L1
phonetic stratal system of L1

Figure 1. *The left shows the linguistic system of a predominately compund bilingual, while the right shows the linguistic system of a predominately subordinate bilingual. Both halves show oral encoding and aural decoding only.*

from the cognitive systems and providing outputs to the articulatory or digitomanual kinetic systems. Each language is a subset of the larger language neurofunctional system (Figures 3 to 9).

1. LATERALITY. In the 1970s and 1980s, experimental psycholinguists were exploring the possibility of less asymmetrical cerebral representation of one or both languages of bilingual speakers. Experimental findings continued to be contradictory in spite of the ever-shrinking relevant population, while there was no clinical evidence whatsoever in favor of any difference of lateralization of language in any type of bilinguals (Paradis 1990, 1992b, 1995). This led me to search for what the increased contribution of the right hemisphere could be, since it obviously was not subserving grammar.

2. THE INTEGRATION OF PRAGMATICS. After considering all possibilities (Paradis, 1987), it appeared that the right hemisphere specialized in processing the pragmatic aspects of language use, which turn out to be at least as important as sentence grammar or the language system per se. The evidence comes from right-hemisphere-damaged patients with dyshyponoia (Paradis 1998). This adds a new component to the model, given that linguistic pragmatic features are language specific. There is thus another dimension to the possibility of unidirectional or bidirectional interference, namely the interpretation of L2 pragmatic clues with their L1 value, hence resulting in misunderstandings and faux pas. What is a polite form of request in one language

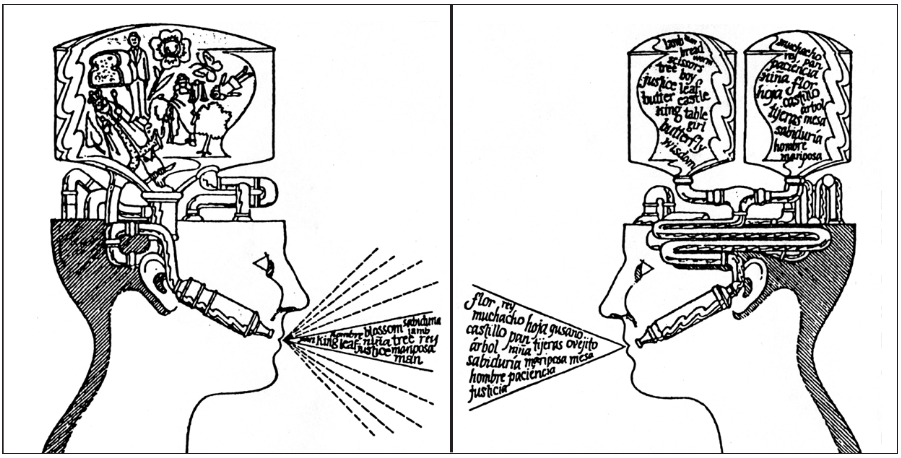
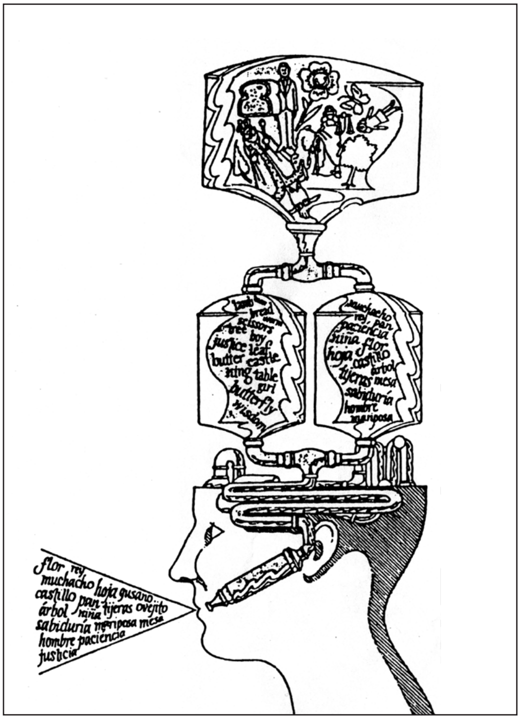


Figure 2a. *Two hypotheses about the way a bilingual person handles information are represented by two arrangements of tanks. One hypothesis (left) is that all his information is stored centrally, or in one tank, and that he has access to it equally with both languages, which are represented by the various taps. The other (right) is that his information is stored in linguistically associated ways, or in separate tanks. Experiments by the author indicated that the actual situation of a bilingual person combines parts of both hypotheses. (Kolers 1968)*

Figure 2b. *The actual situation of a bilingual person may be illustrated by a third hypothesis: both languages are differentially connected to the same conceptual-experiential information store (Paradis 1978).*



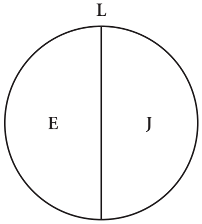


Figure 3. *Schematic representation of the language system (L) comprising 2 subsystems. Left of the vertical bar, English (E); right of the vertical bar, Japanese (J).*

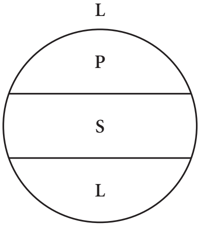


Figure 4. *Language as a collection of modules (neurofunctional subsystems). Only three modules are represented for the sake of simplicity: phonology (P), syntax (S) and lexical semantics (L).*

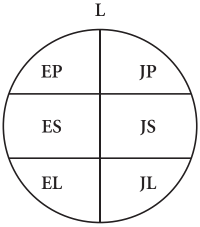


Figure 5. *The languages of a bilingual speaker, each a subsystem within the various language neurofunctional modules. Modules: P, S, L; submodules: EP, JP, ES, JS, EL, JL.*

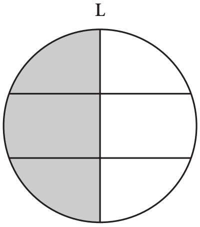


Figure 6. *One language selectively inhibited.*

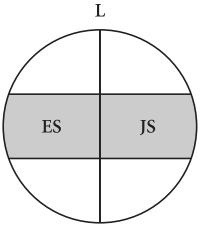


Figure 7. *One module (here syntax) selectively inhibited.*

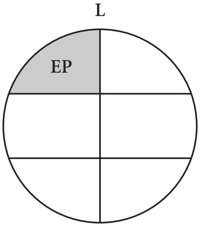


Figure 8. *Selective inhibition of English phonology.*

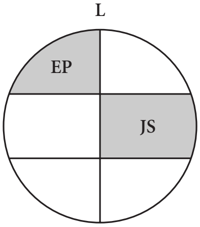


Figure 9. *English phonology and Japanese syntax do not form a natural subsystem. Even though the internal structure of ES is independent of the internal structure of of JS, they both instantiate parameters of S (as opposed to parameters of P, for instance). which makes ES functionally more similar in nature to JS than to EP or JP. A cooccurence of lesions affecting at the same time EP and JS may not be altogether impossible, but it is very improbable, given the different loci subserving the physiology of each neurofunctional submodule. It would take two different lesions (one affecting English phonology, the other Japanese syntax) to give rise to different aphasia, namely, qualitatively different symptoms in each language.*

may be downright rude in another. 'I like your tie' is a compliment in English, but an indirect request for the tie in Chinese or Arabic. 'Why don't you sit down?' is a polite invitation in English, but a rude question in Korean. 'Have a cookie' is a nice thing to say to a guest in English, but considered bossy and hence inappropriate in Navaho. Thus, not only do idiomatic expressions and proverbial phrases have different interpretations, but plain indirect speech acts do too. Not to mention differences in body language, facial expressions and prosody. What counts as a possible indirect speech act, what context determines the use of this or that form of address, conventional metaphors, conventional gestures and facial expressions, all vary in accordance with each language and culture. Thus, in a neurofunctional model of bilingualism, pragmatics also needs to be integrated. Each level of linguistic structure (prosodic, (segmental) phonological, morphological, syntactic and lexical-semantic) is affected by pragmatics during the encoding and decoding of utterances. The listener derives meaning from prosody, the selection of specific words (e.g., *tu/vous,* in French), syntactic constructions (e.g., passive vs. active; topicalized subject or object). A passive construction places emphasis on the patient rather than the agent of the verb and may be reinforced or contradicted by stress at the phonological level, in addition to prosody, facial expression, conventional gestures and the like. The use of *tu*, which will, in one context be interpreted as a sign of friendliness or intimacy, will, by the same hearer in another context, be interpreted as a sign of contempt. There is no available empirical evidence as to whether a bilingual individual can lose access to the linguistic pragmatics of only one language. There is no evidence that would help us decide whether linguistic pragmatics is stored as one extended system or whether, as for the linguistic systems, the pragmatic features associated with each language constitute a subsystem of pragmatics. Our hypothesis is that they do, but this will have to be empirically verified.

Because linguistic pragmatic inference/implicature procedures are specific to each language, we may assume that they form a natural class and that, as such, each is likely to form a subsystem connected to its respective language (subsystem). The schematic representation in Figure 10 (overleaf) is an attempt to visually represent the theoretical construct referred to above.

Pragmatics is an important dimension of bilingualism. Lack of pragmatic knowledge of L2 leads to communication breakdowns at least as important as lack of language (i.e., grammar) proficiency. In the same way that individuals with dyshyponoia encounter social problems, ignorance of L2 pragmatics (especially for speakers with good linguistic skills) is similarly disruptive. Speakers of a language as a second language may understand what native L2 speakers *say*, but not what they *mean*. In turn, they may use inappropriate speech acts or other pragmatic features and thus be misunderstood.

3. CONSCIOUS AND UNCONSCIOUS PROCESSING. For the past century and a half, linguists and aphasiologists have used the term 'language' to mean the language system, the code, the grammar (i.e., phonology, morphology, syntax and semantics), which

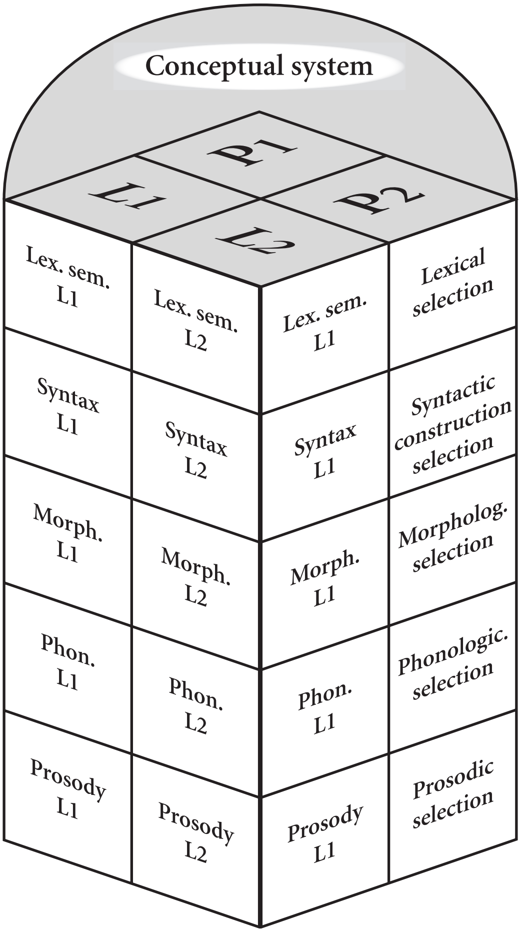


Figure 10. *Visual representation of theoretical constructs described in this paper.*

has come to be referred to as implicit linguistic competence. Language was thus assumed to be represented in circumscribed areas of the perisylvian region of the left hemisphere, since damage to these regions would result in phonological, morphological, syntactic or semantic deficits. Research on second language acquisition pointed to a double dissociation between the ability to do well on off-line metalinguistic tasks (say, on a school test) and on-line accuracy and fluency (say, in normal conversation). This and the double dissociation between aphasia and amnesia drew attention to procedural memory, which underlies linguistic competence, and declarative memory, which underlies metalinguistic knowledge (Paradis 1994). Implicit linguistic competence is acquired implicitly, that is, by focusing attention on something other than what is internalized: children focus on meaning and internalize an underlying syntactic structure that is not there to be observed; they acquire new sounds by focusing on the acoustic properties of the input and end up internalizing the proprioceptive mechanisms that allow them to articulate those sounds, while remaining forever ignorant of which muscle does what, whether the vocal folds vibrate, what the vowel onset time is, and whether their tongue touches the side of their cheeks or the back of the palate. Competence is stored implicitly and remains opaque to introspection. It is used automatically, that is, without conscious control. Metalinguistic knowledge, on the other hand, is learned consciously, stored explicitly (and hence can be recounted), and used in a controlled manner. Competence and knowledge are subserved by different cerebral mechanisms that reside in anatomically different loci. As a result, each is vulnerable to different physiological insult: some bilingual aphasic patients with reduced implicit linguistic competence but available metalinguistic knowledge paradoxically recover a previously weaker L2 over their L1 (Aglioti & Fabbro 1993). Conversely, loss of L2 has been reported in Alzheimer's Disease, a condition in which declarative memory is first affected. (Hyltenstam & Stroud 1993; Obler 1999).

4. MOTIVATION. The extreme variability in second language learners' success has drawn attention to the important dimension of motivation and its cerebral counterpart: the emotive dopaminergic structures and the amygdaloid complex within the limbic system (Schumann 1998). One of the major differences between acquiring a first language and learning a second is that children are highly motivated to understand speech directed at them and to make themselves understood: they speak because they want to convey a message. This first phase of the microgenesis of any utterance is missing in the classroom context where students do not ask 'Where is the pharmacy?' because they have a headache and want to buy an analgesic, but because it is their turn to ask the question that corresponds to the picture on top of page 28 (Paradis 1992a). The reduction in the release of dopamine at the onset of each utterance during the period of appropriation of the language may be an important factor in determining the extent of linguistic competence. Motivation has also been shown to be relevant in differential performance in aphasia. Individuals with global aphasia who are not able to put two words together under normal circumstances may cry out 'Shut the goddam door!' when annoyed at the noise or cold or wind coming from the outside.

5. THE ACTIVATION THRESHOLD. Another component of the model, which is not associated with any particular anatomical area or functional system, but which is physiological in nature, and associated with all higher cognitive representations, whatever their domain, is the Activation Threshold Hypothesis. It proposes that an item is activated when a sufficient amount of positive impulses have reached it. The amount necessary to activate the item constitutes its activation threshold. Every time the item is activated, its threshold is lowered and fewer impulses are required to reactivate it. Thus, after each activation, the threshold is lowered—but it gradually rises again. If the item is not stimulated, it becomes more and more difficult to activate. Attrition

is the result of long-term lack of stimulation. Motivation may have the same effect as repeated practice in lowering the activation threshold.

Pathology (or aging) may disrupt the normal activation levels, causing, for example, word-finding difficulty. Aphasia would correspond to the blanket raising of the threshold for a system or subsystem or module, thus selectively or differentially affecting either the entire language system, or one language, or only the phonology, syntax, or lexical access in one language (see Figures 3 to 9). The hypothesis may account for recency and frequency effects, and priming phenomena, as well as language attrition.

6. NEUROFUNCTIONAL MODULARITY. While the brain is one, and eventually everything is more or less indirectly connected to everything else, some portions of the network are dedicated to specific functions and, as such, are part of different neurofunctional modules. Each module can be selectively impaired by pathology. In the normal use of verbal communication, some modules process information in parallel, while the output of certain ones serves as input to others. Episodic and semantic memory, encyclopaedic knowledge, and information from various contexts serve as the basis for formulating messages which are optionally verbalized, involving at least four independent but collaborating neurofunctional systems (implicit linguistic competence, metalinguistic knowledge, pragmatics and motivation). The extent of reliance on each of these four systems will vary in the bilingual speaker's two languages to the extent that proficiency in the second language differs from that in the native language. While all speakers will primarily rely on implicit linguistic competence to communicate automatically in their native language, they will differ among themselves with respect to their degree of proficiency in L2 and will compensate for the gaps in their linguistic competence by greater but differential involvement of pragmatics and metalinguistic knowledge.

Thus the model presented at the third LACUS Forum 24 years ago has grown in complexity and neurolinguistic relevance but remains valid at the linguistic level of analysis. The subsequent Three-Store and Subsystems Hypotheses, which got their world premiere presentations at subsequent LACUS Forums, have been incorporated into the model, which has now been adopted by and large as the standard working hypothesis in the psycholinguistics and neurolinguistics of bilingualism (see, e.g., De Bot & Schreuder 1993; Green 1998; Grosjean 1998; Malmkiær & Anderson 1991; Pavlenko 1999; Poulisse 1997). The linguistic processes may be sequential (albeit in rapid succession, in milliseconds) or run in parallel. Likewise, during decoding (the interpretation of language input), the pragmatic elements may be processed in succession (i.e., before or after linguistic processing) or in parallel, inferencing from each linguistic component (prosody, word selection, syntactic construction). Encoding may start with pragmatic considerations that guide the selection of linguistic elements, or proceed in parallel, influencing the selection of elements from each level in turn, sequentially. In speaking, activation of the concepts must precede that of the pragmatic and linguistic choices selected to encode the (conceptual) message verbally.

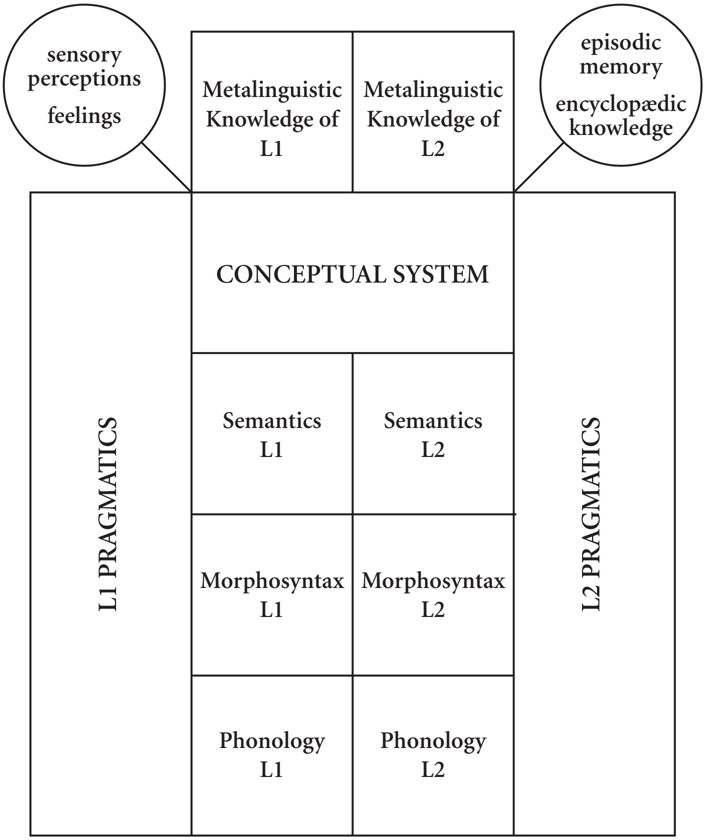


Figure 11. *Schematic representation of the current integrated neurolinguistic theory of bilingualism.*

Figure 11, I guess, is what a schematic representation of the current integrated neurolinguistic theory of bilingualism could look like.

CONCLUSION. The moral of the story is that in a brain where virtually everything is connected to everything else, there are nevertheless portions of the network that are dedicated to the processing of specific functions. While their basic functioning principles are essentially the same throughout the brain, each dedicated portion of the network subserves its own set of representations, the internal structure of which is independent of that of other neurofunctional modules. The performance of a complex task depends on the collaboration of a number of independent neurofunctional modules, each capable of selective impairment. Each language system of a bilingual speaker is a complex subsystem comprising several modules. In addition, in the normal use of language, many other independent neurofunctional modules are necessarily engaged: pragmatic ability, metalinguistic knowledge, concepts, encyclopaedic and episodic memory, as well as motivation. Conceptual mental representations are independent of language and, to the extent that the lexical-semantic constraints of items in each language differ from those of their translation equivalents, language subsystems are connected to some unshared conceptual features (Paradis 2000). There is not a single cerebral mechanism at work in the bilingual speaker that is not also operative to some extent in the unilingual brain. All speakers rely on the same cerebral substrates, albeit to different extents. There may be qualitative differences at the level of *what* is represented but there are only quantitative differences in *how* it is represented and processed.

REFERENCES

Aglioti, Salvatore & Franco Fabbro. 1993. Paradoxical selective recovery in a bilingual aphasic following subcortical lesions. *Neuroreport* 4:1359–62.

De Bot, Kees & Robert Schreuder. 1993. Word production and the bilingual lexicon. In *The bilingual lexicon*, ed. by R. Schreuder & B. Weltens, 191–214. Amsterdam: John Benjamins.

Green, David. W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition* 1:67–82.

Grosjean, François. 1998 Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and cognition* 2:131–49.

Hyltenstam, Kenneth & C. Stroud. 1993. Second language regression in Alzheimer's dementia. In *Progression and regression in language*, ed. by K. Hyltenstam & Å. Viberg, 222–42. Cambridge: Cambridge University Press.

Kolers, Paul A. 1968. Bilingualism and information processing. *Scientific American* 218 (March):78–86.

Malmkiær, K. & J. Anderson. 1991. *The linguistics encyclopedia*. London: Routledge.

Nilipour, Reza & Hassan Ashayeri. 1989. Alternating antagonism between two languages with successive recovery of a third in a trilingual aphasic patient. *Brain and language* 36:23–48.

Obler, Loraine. 1999 Language in dementia: Implications for neurolinguistic theory. In *Language, brain and verbal behavior*, ed. by J. A. Argenter, 63–72. Barcelona: Institut d'Estudis Catalans.

Paradis, Michel. 1977. The stratification of bilingualism. lacus *forum* 3:237–47.

———. 1981. Neurolinguistic organization of a bilingual's two languages. lacus *forum* 7:486–94.

———. 1987. Neurolinguistic perspectives on bilingualism. In *The assessment of bilingual aphasia*, ed. by In Michel Paradis & Gary Libben, 1–17. Hillsdale nj: Lawrence Erlbaum.

———. 1990. Language lateralization in bilinguals: Enough already! *Brain and language* 39:576–88.

———. 1992a. Neurosciences et apprentissage des langues. In *Enseignement des langues dans le monde d'aujourd'hui*, ed. by D. Girard, 21–31. Paris: Hachette.

———. 1992b. The Loch Ness Monster approach to lateralization in bilingual aphasia: A response to Berquier & Ashton. *Brain and language* 43:534–37.

———. 1994. Neurolinguistic aspects of implicit and explicit memory: implications for bilingualism. In *Implicit and explicit learning of second languages*, ed. by N. Ellis, 393–419. London: Academic Press.

———. 1995. Another sighting of differential language laterality in multilinguals, this time in Loch Tok Pisin: Comments on Wuillemin, Richardson & Lynch (1994). *Brain and language* 49:173–86.

———. 1998. The other side of language. *Journal of neurolinguistics* 11:1–10.

———. 2000. Cerebral representation of bilingual concepts. *Bilingualism: Language and cognition* 3:22–24

——— & Marie-Claire Goldblum. 1989. Selective crossed aphasia in one of a trilingual's languages followed by antagonistic recovery, *Brain and language* 36: 62–75.

———, ——— & Raouf Abidi. 1982. Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain and language* 15: 55–69.

Pavlenko, Aneta. 1999. New approaches to concepts in bilingual memory. *Bilingualism: Language and cognition* 2: 209–30.

Poulisse, Nanda. 1997. Language production in bilinguals. In *Tutorials in bilingualism: Psycholinguistic perspectives*, ed. by A.M.B. de Groot & J.F. Kroll, 201–24. Mahwah NJ: Lawrence Erlbaum.

Schumann, John. 1998. The neurobiology of affect in language. *Language learning* 48: Supplement 1.

Weinreich, Uriel. 1953. *Languages in contact*. New York: Publications of the Linguistic Circle of New York.

* * *

WINNER OF THE 2000
PRESIDENTS' POST-DOCTORAL PRIZE

PRESIDENTS' POST-DOCTORAL PRIZE

The Presidents' Post-Doctoral Prize is awarded annually to the paper judged to make the greatest contribution to linguistic knowledge by an author who has completed a doctorate within the preceding ten years, but who does not yet have faculty tenure. The judging panel consists of the current LACUS President and Vice President along with all past presidents in attendance at the meeting.

PARTICLE PLACEMENT IN ENGLISH

Stefan Th. Gries
University of Southern Denmark at Sønderborg

A phenomenon that has attracted considerable interest in linguistics over the last decades is syntactic variation. I will investigate a particular instance of syntactic variation, namely the word order alternation that is possible for English transitive phrasal verbs (TPVs). As an example, consider (1).

(1) a. John picked up the book. construction$_0$ (C$_0$)
b. John picked the book up. construction$_1$ (C$_1$)

This alternation has frequently been referred to as Particle Movement. In this study, I will, as in Gries (1999, 2000) use the term particle placement (PrtPl) in order to avoid the movement metaphor and its theoretical implications.

Over the last 100 years, a large number of studies have been devoted to exploring this type of alternation. More precisely, they have attempted, first, to provide an adequate structural description of the two possible constituent orders and, second, to find the variables that determine native speakers' choices governing the alternation[1].

Interestingly, a (at a superficial glance) simple word order alternation seems to be influenced by a large number of variables from many subdisciplines of linguistics. Table 1 (overleaf) gives an overview of variables that have been claimed to contribute to PrtPl. The middle column names the variable whereas the other columns provide the values/levels purportedly correlating with a preference for a construction.

However comprehensive this list of variables may seem at first sight, there are three methodological issues I would like to address (for a conceptual critique of some of these variables, cf. Gries 2000).

First, nearly all of the analyses have been monofactorial in nature. That is, most scholars have concentrated on one variable at a time and have attempted to support its assumed effect on PrtPl. Note that this does not mean they have not acknowledged that several variables are important; it only means that each variable's effect has been investigated in isolation without concern for the simultaneous effect of other variables. Let us consider an example. Fraser (1974:571) has argued that verbs without initial stress entering into TPVs are preferred in construction$_1$. In order to substantiate his claim, he provides the following minimal pairs and acceptability judgements:

(2) a. ?I will insult back the man.
b. I will insult the man back.

Value/Level for C_0	**Variable**	**Value/Level for C_1**
stressed DO	stress pattern of VP	stressed particle
	phonetic shape of V	no initial stress
	NP type of DO	(semi-)pronominal
definite	determiner of DO	indefinite / none
long	length of the DO	
complex	complexity of the DO	
idiomatic	meaning of the VP_1	
inanimate	animacy of the DO[2]	animate
abstract	concreteness of the DO	concrete
low	entrenchment of the DO	high
long	distance to last mention of the DO	short
low	times of preceding mention of the DO	high
low	cohesiveness of the DO to the preceding discourse	high
short	distance to next mention of the DO	long
high	times of subsequent mention of the DO	low
high	cohesiveness of the DO to the subsequent discourse	low
	following directional adverbials	yes
high	production difficulty	low

Table 1. *Variables governing PrtPl (as suggested in the literature)*[3].

(3) a. ?We converted over the heating to steam.
    b. We converted the heating over to steam.

(4) a. ?They attached up the tag on the wall.
    b. They attached the tag up on the wall.

However, the question arises as to how Fraser knows that it is the phonetic form of the verb that is responsible for the purported preference of the (b) sentences rather than some other variables he also discusses in the same paper. The indicated preference of construction$_1$ in all three examples (if it is representative at all) could be due to the DO being simple, very short and definite rather than Fraser's phonological variable. Obviously, the data Fraser cites simply do not warrant his conclusion as they fail to take into consideration these additional variables. This instance shows that, in complex cases like PrtPl, minimal pairs can distort the picture more than they are helpful in spite of their time-honoured place in linguistics.

What is more, it is uncontroversial that, in complex phenomena (given the large number of variables mentioned above), independent variables can interact with each other with respect to the impact they have on the dependent variable. Such phenomena can hardly be identified by using classical minimal pair tests, which contributes further to the lack of an adequate description of PrtPl. Therefore, I advocate the use of multifactorial techniques, which enable the researcher to describe PrtPl more efficiently.

Moreover, with few notable exceptions (Chen 1986; Gries 1999, 2000; Hawkins 1991, 1994) analyses of PrtPl have been based on intuitive and introspective analyses of grammaticality/acceptability judgements alone. This methodology may be found acceptable (and even rewarding) in particular research traditions (most notably transformational-generative grammar). From my point of view, however, this methodology is fundamentally flawed. To name one reason for this opinion, numerous publications have shown (cf., e.g., Labov 1975) that relying on grammaticality/acceptability judgements alone does not meet well-established standards of scientific research (objectivity, validity, and reliability). Thus, my analysis will rely solely on naturally-occurring instances of PrtPl (from the British National Corpus).

Second, in spite of many decades of research, there is still no account of PrtPl that tries to explain why speakers choose one construction over the other in a particular discourse situation. Such an account would, first, identify all relevant variables and subsume them under a (probably) small set of factors and, second, identify all irrelevant variables on a principled basis and eliminate them from further consideration. There are several such analyses which have attempted to provide a unifying hypothesis for the main motivation of the alternation by relating PrtPl to issues of givenness/topicality (which boils down to degree of activation) of the DO's referent (Erades 1961; Chen 1986) or to purely syntactically conditioned processing effort (Hawkins 1991, 1994). However, these previous investigations do not go far enough for two reasons, as follows.

On the one hand, there are only two major empirical analyses of PrtPl (Chen 1986, and Hawkins 1991, 1994), and both suffer from several drawbacks. (The following discussion will only deal with two problems. For a more comprehensive critique, cf. Gries 2000). Both Chen and Hawkins operationalise the givenness of the DO's referent by counting co-referential expressions in the preceding context, a methodology advocated in numerous papers by Givón (cf, e.g., Givón 1983, 1988). However, it has been shown beyond any reasonable doubt that the referent X of some expression is not only activated by strictly co-referential items. Rather, the inferential processes on the part of hearer in particular discourse situations also activate related concepts. While the notion of *related concept* can be criticised as being fairly vague, studies such as Clark (1977) on bridging and by Bolkestein (1985) as well as Bolkestein and Risselada (1987) on cohesiveness have shown that the referent X of some expression can in fact (depending on the particular characteristics of the discourse) be activated by hyperonyms of X, hyponyms of X, co-members of X's level of categorisation, parts of X, the function of X, reasons leading to X, consequences following from X, etc. Therefore, analyses of the influence of givenness/topicality had better incorporate these variables if the investigation of givenness is to be valid. What is more, both the analyses by Chen and Hawkins have been carried out with little degree of statistical sophistication as, e.g., neither analysis uses any significance tests to support their far-reaching conclusions.

On the other hand, Chen's and Hawkins' proposals also fail to account for a variety of factors which cannot easily, if at all, be related to givenness/topicality or purely syntactically conditioned processing effort, namely the idiomaticity of the VP, VP-final directional adverbials, and the concreteness of the DO, to name a few. Due to these and other drawbacks, an explanatory account of PrtPl that is both coherent and empirically supported cannot be found.

Third, given the shortcomings just discussed, it comes as no surprise that it has not been possible to predict which construction a native speaker will choose. This is a problem that derives from the lack of truly multifactorial treatments of the subject matter.

In the remainder of this paper, I will propose a hypothesis that explains the constructional choices of speakers given a particular discourse context (cf. section 2) and I will show how, by means of using multifactorial statistics, this hypothesis can be supported and native speakers' choices of construction can be predicted with a high degree of accuracy (cf. section 3).

1. THE PROCESSING HYPOTHESIS. In this section, I propose a hypothesis to explain the distribution of the two constructions and outline which variables are, according to this hypothesis, relevant and which are not.

> **The Processing Hypothesis (PH)**: By choosing one of the two constructions for an utterance U a speaker S formulates U in such a way as to communicate the intended message with as little processing effort as possible. More specifically, for most variables at least, this means that construction$_0$ will be preferred for verb-particle constructions (VPCs) with DOs requiring a lot of processing effort; construction$_1$ will be preferred for VPCs with DOs requiring little processing effort[4].

In order to investigate this hypothesis, two different steps are necessary. On the one hand, it must be shown how the above-mentioned variables relate to processing effort. This will be done in the remainder of this section. On the other hand, it must be shown empirically, that (i) the variables included by the PH correlate with PrtPl in the direction predicted above and (ii) the variables excluded by the PH do not correlate with or contribute to PrtPl; this will be the topic of section 2.

1.1. PHONOLOGICAL VARIABLES. The stress pattern of the VP can be straightforwardly related to processing requirements: in functional analyses of information structure, it has been useful to distinguish two kinds of information, namely given and new information. Generally, in English the new (and/or important) information is positioned sentence-finally. Moreover, it is by now common ground that stress on a linguistic expression typically serves to indicate the newness or importance of the referent of this linguistic expression, i.e., to direct the attention of the hearer to the respective referent, thereby increasing the processing cost associated with that referent. Combining these two findings results in the distribution predicted in the PH: the expressions whose referents are to be processed thoroughly occur clause-finally and are stressed.

1.2. MORPHOSYNTACTIC VARIABLES. As far as the TPV (transitive phrasal verbs) is concerned, construction$_0$ facilitates processing because the speaker can complete (and thus disambiguate) the TPV earlier if the particle immediately follows the verb and need not be borne in mind until the DO has been uttered completely (cf. Rohdenburg 1996). But what about the DOs of TPVs? Let us start with the variable *NP type of the DO*.

This variable is, in fact, quite closely connected to utterance processing: personal pronouns and semi-pronominal/referentially vague nouns are only used when their referents are identifiable and active, whereas lexical noun phrases are much more likely to be used with unused and brand-new referents. Again, the distribution is as predicted in the PH: active referents of personal pronouns require little processing effort and correlate with construction$_1$, referents of lexical DOs are, on the whole, more likely to require more processing and occur preferably in construction$_0$.

Likewise, the variable *determiner of the DO* noun phrase is also concerned with processing aspects:

> Linguists traditionally deal with the binary distinction between definite and indefinite, with the former marking topics which the speaker assumes the hearer can identify uniquely, is familiar with, are within his file (or register) and thus available for quick retrieval. On the other hand, indefinites are presumably topics introduced by the speaker for the first time, with which the hearer is not familiar, which therefore are not available to the hearer readily in his file. (Givón 1983:9f.)

Comment is hardly called for: definite determiners are said to prevail in construction$_1$ and indefinite determiners (used for unused or even brand-new referents requiring conscious activation) are said to prevail in construction$_0$ so that the pattern found is as expected.

Length of the DO and *complexity of the DO* (irrespective of how these are measured; cf. Wasow 1997) can be dealt with simultaneously. Obviously, longer/more complex NPs require more processing effort while shorter/less complex noun phrases can be more easily processed. But apart from this purely structurally motivated approach, there is also a functional principle at work: ‘the new information often needs to be stated more fully than the given (that is, with a longer, “heavier” structure)’ (Quirk et al. 1985:1361). Thus, if the newness of a referent, on average, renders DOs long and complex (in order to provide the necessary information for the hearer to establish a new referential file), the larger amount of linguistic material requires more processing effort than the one needed for given information. This, in turn, increases the likelihood that production and planning effects or mistakes such as hesitations, repetitions, and false starts can be observed in actual speech (cf. Arnold & Wasow 1996). Ultimately, both the length and its functional motivation go hand in hand so that again much processing effort is linked to construction$_0$ and little processing effort to construction$_1$.

1.3. SEMANTIC VARIABLES. We know that VPCs come in two different word orders, and we also know that the meanings of VPCs range from literal to idiomatic. In construction$_1$, the particle is positioned in the canonical clause-final position for focal elements, so that the particle is processed more intensively than the DO. Thus, the word order of construction$_1$ naturally underscores the spatial contribution the particle makes to the compositional meaning of the TPV in the utterance. Therefore, construction$_1$ is the natural choice for a speaker communicating a state of affairs where the spatial meaning is prominent[5].

Very often in the case of idiomatic constructions, the meaning of the TPV is not compositional. The particle does not just add some spatial information to a straightforward sense of the verb (as with literal TPVs), but the idiomatic TPV conveys a meaning that is not a function of the meaning of its parts and their interrelations, but must be stored on its own. In other words, when the speaker accesses the meaning of the TPV for production, then the complete idiomatic TPV (i.e., verb *and* particle) are simultaneously accessed. Thus it is only natural that the verb and the particle are uttered following one another directly: it would be uneconomical to process the opaque meaning of a TPV but produce the parts that trigger this opaque meaning in possibly widely disparate positions of the sentence[6].

Turning to the next variable, no significant effect of *animacy of the referent of the DO* on particle placement is to be expected according to the PH: there is no reason to assume that animate referents yield context-dependent processing requirements substantially different from inanimate referents, and there is also no reason to assume that animate referents are more likely to undergo caused motion. Therefore, I will investigate animacy but consider it not to contribute to the constructional alternation.

Finally, the variable *entrenchment of the DO* (as analysed in Gries 1999) will not be analysed directly because the entrenchment hierarchies offered in Deane (1992) and Gries (1999) comprise several variables that will be dealt with separately (and thus much more precisely) in this analysis.

1.4. DISCOURSE-FUNCTIONAL VARIABLES. The discourse-functional variables concerning the preceding context relate straightforwardly to matters of processing effort. Information that is given as a result of having (frequently and/or recently) been evoked in the preceding discourse or being readily inferable from the preceding context requires less processing effort than discourse-new or especially hearer-new information. Thus, the distribution of the two constructions (and their processing cost) relative to these variables is the one given above in Table 1. However, the discourse-functional variables concerning a subsequent context cannot be related to the processing cost of the VPC. Even if it were true that these variables measure the importance of the DO's referent, it is difficult to see how a speaker should be able to foresee precisely the development of the discourse to follow (cf. also Hawkins 1994:225). Thus, I claim that these variables will neither be relevant to nor correlate with the processing cost of the utterance.

	Spoken	Written	Row Totals
Construction$_0$	67	127	194
Construction$_1$	133	76	209
Column Totals	200	203	403

Table 2. *Corpus data*[7].

1.5. OTHER VARIABLES. The next variable is concerned with the presence of a directional adverbial following either the verb or the particle. If a directional adverbial follows the VPC, then it typically serves to elaborate either the path along which the referent of the DO is being moved, (5)a, or the resultant location of the referent of the DO, (5)b and endnote 5.

(5) a. So Tom took Peter along past the new Pump House.
b. Fred put the book down on the table.

For construction$_1$, where the spatial meaning is foregrounded (cf. above section 2.3), it is therefore quite natural to expect additional material (in the form of a directional adverbial) providing additional information on the direction or the endpoint of the movement process. On the other hand, construction 0 does not normally denote a movement process that can be further elaborated with information concerning directionality so following directional adverbials are, though not strictly ruled out, not to be expected. Thus, the distribution predicted in the PH seems to be fully justified.

2. THE STATISTICAL INVESTIGATION: DATA AND RESULTS. In order to empirically test the PH conforming more to the standards mentioned in section 1, I studied a set of 403 VPCs from the BNC. Table 2 gives an overview over the sentences that were analysed.

By analysing the 403 sentences and their context (the 10 preceding and 10 subsequent clauses)[8], each VPC was assigned values representing the values/levels of the variables listed in Table 1[9]. Then, two different procedures were used: first, for each variable, a strictly monofactorial coefficient of correlation was calculated (section 2.1). This may seem surprising, given that I have devoted some space to arguing that monofactorial analyses are ill-suited for problems of such a level of complexity. The reason for including monofactorial correlations is that I wanted to subject all variables to an empirical test, as few variables have already been tested on naturally-occurring data. Second, a multifactorial analysis was conducted in order, first, to find out how much variance of PrtPl we can account for (given the variables we know of) and, second to try to predict native speakers' choices of a construction in natural discourse (section 2.2).

2.1. MONOFACTORIAL RESULTS. Due to the different levels of the variables within the analyses, a variety of different coefficients of correlation had to be computed. To cut a long story short, Table 3 contains the correlation coefficients for each variable.

Variables that were analysed empirically	**Correlation coefficient**
Complexity of the DO (complex)	$\gamma=-0.847$ ***
Idiomaticity of the verb phrase (idiomaticity)	$\gamma=-0.6$ ***
Length of the DO in syllables (LengthS)	$r_{pbis}=-0.504$ ***
NP Type of the DO (type)	$\phi=0.492$ ***
Distance to last mention of the DO (ActPC)	$r_{pbis}=0.452$ ***
Length of the DO in words (LengthW)	$r_{pbis}=-0.45$ ***
DO's cohesiveness to the preceding discourse (CohPC)	$r_{pbis}=0.429$ ***
Times of preceding mention of the DO (topm)	$r_{pbis}=0.415$ ***
Last mention of the DO (lm)	$\phi=0.411$ ***
Overall mention of the DO (om)	$r_{pbis}=0.358$ ***
Concreteness of the DO (concrete)	$\phi=0.339$ ***
Determiner of the DO (det)	$\phi=0.319$ ***
Directional adverbial following the DO (PP)	$\phi=0.268$ ***
Times of subsequent mention of the DO (tosm)	$r_{pbis}=0.192$ ***
Animacy of the DO (animacy)	$\phi=0.166$ ***
DO's cohesiveness to the subsequent discourse (CohSC)	$r_{pbis}=0.143$ **
Next mention of the DO (Nm)	$\phi=0.104$ *
Distance to next mention of the DO (ClusSC)	$r_{pbis}=0.1$ *
Production and planning effects (disfluency)	$r_{pbis}=-0.006$ ns

Table 3. *Monofactorial results of the analysis.*

Generally speaking, the higher the absolute value of the correlation coefficient, the more important this variable is for PrtPl. But before we begin to interpret these results, a comment concerning the significant correlation between the variable animacy and PrtPl is necessary. This result might be taken as evidence that the PH is not fully correct since I predicted that animacy should not play a role. This, however, would be mistaken as the variable animacy not only tells us whether the referent of the DO is animate or not; it also tells us something about the concreteness of the DO: if the referent is animate, it must be concrete. Thus, if we want to test the influence of Animacy alone, we need to take out the influence of concrete on animacy. There is a statistical technique serving this purpose, namely that of partial correlations, and if we do that, then Animacy does not have any influence on PrtPl anymore: $r_{\text{animacy construction} \bullet \text{concrete}} = 0.04$; $t_{400} = 0.757$; $p = 0.449$ ns. In less technical terms: animacy only had a significant effect because it tells us something about the concreteness of the DO's referent: if we take out that information, animacy turns out to be irrelevant while concrete remains important (as predicted by the PH). This is a good example of how monofactorial analyses can result in 'statistical evidence' that, on closer inspection, turns out to result from the artificial separation of variables that are actually closely related.

While it is not possible to go through all of the results in detail, some important conclusions can be drawn from the monofactorial analysis. First, we find a rough ranking of variable groups: morphosyntactic variables are the most important ones, followed by semantic variables and discourse-functional variables pertaining to the preceding context; least important are the discourse-functional variables concerned with the subsequent discourse. Second, on the whole, the PH receives strong support: the variables included by the PH all correlate with PrtPl, one, significantly and, two, in the predicted direction. Third, there are some variables (TOSM, CohSC, NM and ClusSC) which correlate significantly with PrtPl although the PH has predicted otherwise (interestingly the correlation is exactly the opposite one than that observed by Chen). In this respect, the PH is not supported, but we will return to these variables below. Finally, we have, for the first time, a clue as to the strength of the individual variables in isolation, a level of precision not attained so far.

The following section, however, will approach PrtPl in yet a different way.

2.2. MULTIFACTORIAL RESULTS. While the results of section 2.1 are already a major leap forward (given the absence of similarly thorough analyses so far), they are still not quite satisfactory. The problem is, as was already mentioned, that monofactorial analyses do not enable us to address the issue rewardingly. Consider the choice of construction from the perspective of speech production: Obviously, no given native speaker computes a series of monofactorial correlations or weighs variables independently in order to reach a decision as to the construction to be uttered (although the method of monofactorial analyses seems to imply just that). Rather, for the native speaker, all variables' values/levels are present at the point of time where the constructional choice takes place. Thus, any cognitively realistic account needs to incorporate all the variables simultaneously rather than artificially isolated.

Let us first try to measure the overall success of the research on PrtPl in a multifactorial way. Using the General Linear Model (a generalisation from techniques such as regression analysis and analysis of variance), we can assess the amount of variance of the dependent variable PrtPl that can be explained both by all variables ever postulated and, more interestingly for our purpose, for the PH. Consider Table 4 (overleaf)[10].

In sum, the adjusted multiple correlation coefficient R_c for the variables included in the PH is not only highly significant; it is even slightly higher than R_c for all variables ever postulated to influence the alternation. In less technical terms, if we include variables into the analysis other than those required by the PH, what we get is random noise. Thus the PH receives in this case overwhelming support. However, we would still like to know exactly how the variables behave and how we can predict speakers' choices.

With a so-called discriminant analysis, we can enter all variables into an equation and figure out how important each variable is (by means of a factor loadings) and what choice of construction the analysis would predict. Table 5 provides the results

Statistical index	All variables	Variables of the PH
R	0.79	0.761
R_c	0.719	0.723
F	$F_{94, 308}=4.4$	$F_{53, 349}=9.074$
p	<0.001 ***	

Table 4. *Multifactorial results of the analysis.*

Variables of the discriminant analysis	Loading	Effect on PrtPl
Length of the DO in syllables	-0.545	high variable values → construction$_0$ low variable values → construction$_1$
Lexical DO	-0.496	
Intermediate complexity of the DO	-0.477	
Length of the DO in words	-0.47	
Idiomatic VP	-0.323	
Indefinite determiner of the DO	-0.281	
Production/Planning difficulty	-0.006	given the low factor loadings (fl) (-0.22<fl<0.22), these variables do not discriminate well between the two constructions
Semipronominal DO	0.086	
Distance to next mention of the DO	0.093	
Is there a next mention of the DO?	0.097	
DO's cohesiveness to the subs. disc.	0.134	
Animacy of the DO	0.157	
Times of subs. mention of the DO	0.183	
No determiner of the DO	0.222	high variable values → construction$_1$ low variable values → construction$_0$
VP-final directional adverbial	0.277	
Literal VP	0.308	
Concreteness of the DO	0.336	
Is there a last mention of the DO?	0.42	
Times of prec. mention of the DO	0.426	
DO's cohesiveness to the prec. disc.	0.443	
Distance to last mention of the DO	0.473	
Pronominal DO	0.494	
Simple DO	0.571	
Wilks' Lambda=0.465; canonical correlation=0.73; p<0.001 ***		

Table 5. *Results of the discriminant analysis for all variables.*

Composition of the learning sample	Composition of the prediction sample	Correct predictions
200 oral + 150 written sentences	53 written sentences	81.1%
150 oral + 200 written sentences	53 oral sentences	67.9%
174 oral + 176 written sentences	26 oral + 27 written sentences	88.7%

Table 6. *Prediction accuracy of three analyses.*

of a discriminant analysis for all variables (nominal and ordinal variables were 0, 1-coded; variables failing the tolerance test were omitted).

Given space limitations, three findings deserve attention: first, again all variables included into the PH correlate with PrtPl strongly and in the predicted direction. Second, the overall ranking of the variable groups found in the monofactorial analysis is preserved. Lastly, the most realistic multifactorial analysis further supports the PH in that all the variables that I have argued to be irrelevant (namely ANIMACY and all the discourse-functional variables concerned with the following context) do indeed get only very low factor loadings.

What is more, when we perform a second discriminant analysis including only the variables of the PH, then we can test how accurate the predictions of constructional choices from the PH really are. If we enter all 403 cases into the analysis, the discriminant analysis computes a constructional choice, and we can compare the accuracy of this classification with the 'real' choices by native speakers in their discourse settings. The classification accuracy resulting from this procedure is very high: 85.9% of all the 403 utterance are classified correctly. However, this is not really prediction since the cases to be classified also figured in the analysis from which the discriminant function was derived. Therefore, I have pursued two ways to further support the argument of the predictive power of the PH.

First, I split the sample into two parts, one consisting of 350 cases, the other of 53 cases. The former sample was a learning sample to which I applied a discriminant analysis to obtain a discriminant function; the latter sample was a prediction sample whose constructions were predicted on the basis of the discriminant function derived from the learning sample. In order to anticipate criticism of my possibly biased choice of the samples, I performed this test three times with different learning and prediction samples. Table 6 shows the composition of the samples and the results.

Obviously, the cross-validation also strongly supports the PH: on the whole the constructional choices in the prediction sample can be predicted quite accurately on the basis of the learning sample (with oral data being most difficult to predict, which is, I think, due to the more spontaneous and interactive nature of oral speech as compared to planned writing).

However, the ideal way of testing the predictive power of models is via cross-validation, using the leave-one-out method (also called U-method). If we apply this procedure to the present data, we perform 403 analyses in each of which we predict the choice of construction in a single case on the basis of the remaining 402 cases. This again guarantees that no case is used for its own prediction. The result of this cross-validation for the present analysis is a prediction accuracy of 83.9%. This result is for all practical purposes very much the same as the one we had for the classification accuracy and the split-sample technique, which shows that the results are quite robust and the predictive power of all the variables together is indeed exceptionally high.

3. DISCUSSION/CONCLUSION. We have seen that the Processing Hypothesis subsumes more variables under a single explanatory account than any other previous hypoth-

esis. While the present analysis is not the first one to relate PrtPl to processing cost (cf. Hawkins 1994), it is important to notice that it is the first one that integrates all previous variables rather than excluding them *a priori*. We have also seen that the PH excludes several variables from further consideration. Both of these predictions are borne out by the data to such an extent that we can now, on the basis of an argumentatively sound hypothesis correctly predict 84% of native speakers' choices for a particular construction in a natural discourse setting.

One question, however, remains unanswered. The PH was neutral with respect to whose processing effort we were concerned with although it would be preferable to be able to explicitly relate the processing cost associated to the variables under investigation to one or both of the interlocutors. However, as Arnold et al. (2000) have shown, this is hard to achieve since what is beneficial to the speaker is very often beneficial to the hearer, too. To name just one example, long complex DOs require more processing for the speaker, so he might prefer it to process them once everything else has already been processed. At the same time, hearers also prefer the sentence-final position for complex DOs in order to reduce their memory load in parsing. Thus, the question of shoes processing cost we are primarily concerned with awaits further (experimental) research.

More importantly, however, are probably the methodological ramifications. The study has shown that multifactorial analyses of naturally-occurring data can go way beyond previous kinds of analyses. Not only are these techniques more likely to yield cognitively realistic results, they also enable us to identify spurious correlations (i.e., statistical artefacts, recall the variable ANIMACY) and, last but not least, compare the predictive power of competing analyses. Suppose that some researcher claims to have found additional variables influencing PrtPl or a hypothesis making different predictions. In that case, we simply analyse a set of naturally-occurring sentences with respect to all variables (his/hers and mine) again and test which hypothesis results in a higher prediction accuracy. In the absence of additional indications to the contrary, the analysis with a higher degree of predictive power is to be preferred. In conclusion, I hope to have also shown how rewarding a truly multifactorial perspective on the analysis of syntactic variation can be.

---

[1] Here as well as in the remainder of the paper, the expressions *choice of construction* or *speakers' decisions* are not meant to imply that there is necessarily a conscious choice on the part of the speaker.

[2] Of course, animacy and concreteness of the DO refer to the animacy and the concreteness of the referent of the DO.

[3] It is also interesting to note in passing that nearly all of these variables have already been identified single-handedly by Van Dongen (1919). Put differently, not much progress has been made since then as, unfortunately, many analyses have failed to notice this and other early grammarians' insightful works.

[4] This formulation leaves open the question of whose processing effort we are concerned with. This question is not relevant to the subsequent discussions, but it will be addressed briefly in section 3.

[5] Construction$_1$ apparently instantiates a subtype of the caused-motion construction in Construction Grammar (cf. Goldberg 1995 chapters 3 and 8), which has the following basic meaning: 'the causer argument directly causes the theme argument to move along a path designated by the directional phrase; that is, X CAUSES Y TO MOVE Z' (Goldberg 1995:152). Similarly, in TPVs like, say, *Fred brought the book back*, the causer (*Fred*) directly causes the theme argument (*the book*) to move along a path or up to a point designated by the directional phrase (*back*). For this reason, it also follows that concrete referents of DOs will preferably occur in construction$_1$ while abstract objects probably will not: concrete referents can undergo caused motion whereas abstract referents cannot. Finally, note that Goldberg's caused-motion construction is metaphorically related to the resultative construction, which fits previous observations that TPVs can also license a resultative meaning as well.

[6] This claim is supported by the independent observation that idiomatic expressions are in general much less susceptible to syntactic rearrangements than literal expressions.

[7] The majority of the VPCs investigated consist of the most frequent verbs and particles entering into VPCs. The required frequency data in turn result from my own collection of 1,357 different TPVs.

The question may arise why 'only' 403 examples were used for the analysis. First, it has to be observed that, with 403 cases, this is by far the largest quantitative corpus-based analysis of particle placement ever reported (cf. Hawkins' (1994) analysis of a mere 179 cases or Chen's (1986) analysis of only 239 cases). Additionally, the results will show that the predicted effects are all quite strong and highly significant, supporting my claim that the number of cases is in fact not too small at all.

[8] In order not to overly minimise the context, I did not count questions tags, discourse markers (e.g. *you know, I mean*) and repetitions / false starts.

[9] The variable semantic modification of the particle (with words like *right*) was not investigated because the corpus data did not contain a single example of such cases. The influence of stress was not investigated since the available corpus data were not phonologically annotated.

[10] The analysis included two- and three-way-interactions; full-factorial designs of this complexity are difficult to calculate and even more difficult to interpret.

REFERENCES

Arnold, Jennifer E. & Thomas Wasow. 1996. Production constraints on particle movement and dative alternation. Poster presented at the CUNY Conference on Human Sentence Processing.

Arnold, Jennifer E., Thomas Wasow, Anthony Losongco & Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76:28–55.

Bock, J. Kathryn. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological review* 89:1–47.

Bolkestein, A. Machtelt. 1985. Cohesiveness and syntactic variation: Quantitative vs. qualitative grammar. In *Syntax and pragmatics in functional grammar*, ed. by A. Machtelt Bolkestein, Casper de Groot & J. Lachlan Mackenzie, 1–15. Dordrecht: Foris.

——— & Rodie Risselada. 1987. The pragmatic motivation of syntactic and semantic perspective. In *The pragmatic perspective: Selected papers from the 1985 International Pragmatics Conference*, ed. by Jef Verschueren & Marcella Bertucelli-Papi, 497–512. Amsterdam: John Benjamins.

Chen, Ping. 1986. Discourse and particle movement in English. *Studies in language* 10:79–95.

Deane, Paul D. 1992. *Grammar in mind and brain: Explorations in cognitive syntax*. Berlin: Mouton de Gruyter.

Erades, Peter A. 1961. Points of modern English syntax. *English studies* 42:56–60.

Fraser, Bruce. 1974. Review of *The phrasal verb in English* by Dwight Bolinger, *Language* 50:568–75.

Givón, Talmy (ed.). 1983. *Topic continuity in discourse: A quantitative cross language study*. Amsterdam: John Benjamins.

———. 1988. The pragmatics of word order: Predictability, importance, and attention. In *Studies in syntactic typology*, ed. by M. Hammond, Edith A. Moravcsik & J. Wirth, 243–84. Amsterdam: John Benjamins.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gries, Stefan Th. 1999. Particle movement: A cognitive and functional approach. *Cognitive linguistics* 10:105–46.

———. 2000. *Towards multifactorial analyses of syntactic variation: The case of particle placement*. PhD Dissertation, University of Hamburg.

Hawkins, John A. 1991. Syntactic weight versus information structure in word order variation. In *Informationsstruktur und Grammatik*, ed. by Joachim Jacobs, 196–220. Opladen: Westdeutscher Verlag.

———. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Kruisinga, E. & Peter A. Erades. 1953. *An English grammar*, vol. I. Groningen: P. Noordhoff.

Labov, William. 1975. Empirical foundations of linguistic theory. In *The scope of American linguistics*, ed. by Robert Austerlitz, 77–133. Lisse: Peter de Ridder Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language.* London: Longman.

Van Dongen, W. A. Sr. 1919. *He puts on his hat* and *He puts his hat on*. *Neophilologus* 4:322–53.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language variation and change* 9:81–105.

WINNER OF THE 2000

PRESIDENTS' PRE-DOCTORAL PRIZE

PRESIDENTS' PRE-DOCTORAL PRIZE

The Presidents' Pre-Doctoral Prize is awarded annually to the paper judged to make the greatest contribution to linguistic knowledge presented by an author who has not yet completed a doctorate. The judging panel consists of the current LACUS President and Vice President along with all past presidents in attendance at the meeting.

NOBILITY IN THE HISTORY OF ADJECTIVES ENDING IN *-IC* AND *-ICAL*

Mark Kaunisto
University of Tampere

INTRODUCTION. In the study of word-formation, the possibility of using large text corpora has in the last few years opened whole new avenues of research. Whereas earlier studies have pointed out and explained different kinds of phenomena and types of word-formation, it has been somewhat difficult to present in an exact and detailed manner the extent and proportion of e.g. changes in a language, let alone to prove the validity of some of the propositions made in this field (sometimes possibly based on intuition plus only a few examples). The historical aspects in some of the basic textbooks of word-formation, for example, Marchand's (1969), are heavily based on findings from the *Oxford English Dictionary*, and at the time of the book's release, Marchand's perceptive observations undoubtedly seemed to have a very solid base. Now that we have the ability to do more computerized searches from large corpora and thus to focus on even the smallest units in detail, the future will no doubt see, to some extent, a more in-depth rewriting of the basics of English word-formation and, as Kennedy (1998:6) has pointed out, altogether new theoretical generalisations.

From the viewpoint of corpus linguistics, word-formation seems to be a field where the topics of research may set rather high requirements for corpus size in order to achieve scientifically valid findings. What is a suitably large corpus depends, of course, very much on what one is investigating. Some word-formation types (e.g. blending) can be difficult to examine because of the relative infrequency of such words in almost any collection of texts. Suffixation in general is an issue that has been examined from different aspects (see e.g., Biber et al. 1998:63), but studies on the use of individual suffixes are fewer. In this paper I will outline earlier studies on the variation between adjectives ending in *-ic* and *-ical,* and introduce my current work on the topic, with special attention given on their use in the Early Modern English period.

1. THE SUFFIXES *-IC* AND *-ICAL*. The incidence of the suffixes *-ic* and *-ical* presents a complex picture in the vocabulary of present-day English, and there are cases in which both suffixes are used in connection with the same root, as in *analytic/analytical* and *philosophic/philosophical.* The existence of such competing adjective pairs has been observed by several grammarians, one of the first being James Elphinston in 1765, who comments on some differences between the general semantic content of the two endings. The issue has been looked into by late 19th century grammarians and 20th century lexicographers, but at closer examination it becomes evident that the question of when each variant should be used is unanswered. The aim of my research is to study the ways in which such adjectives have been used in different times, and how their use has

changed, describing the different currents of change and the linguistic factors involved (syntactic, semantic, etc.) and also to see why such changes took place.

The models for the English suffixes *-ic* and *-ical* were the French words ending in *-ique*, Latin words ending in *-icus* and *-icalis*, and Greek words ending in *-ikós*. The first words with the suffix *-ic* appeared in Old English, and the *-ical* formations appear in the Middle English period. The present-day situation of free variation between some adjectives in *-ic* and *-ical* (e.g., in the British National Corpus, the numbers of occurrences for *pedagogic* and *pedagogical* are roughly equal, and there seem to be no apparent differences as to their meaning or collocations) is a result of the fact that in the Early Modern English period, both endings were used to produce new adjectives, the functions of the endings being almost identical ('of, belonging or pertaining to, resembling X').

Earlier accounts on the developments in the use of *-ic* and *-ical* may seem somewhat confusing: for example, some of Marchand's comments suggest that forms in *-ical* would be in general more common than the *-ic* forms, and in the case of competing pairs, the longer forms are 'stronger'. He also states that, as a result of competition between two synonymous adjectives in *-ic* and *-ical*, the shorter forms are usually the ones that have been thrown out (1969:242). These comments are in opposition to Earle's (1892:404) observations, according to which the appearance of both nouns and adjectives in *-ic* caused 'a demand for an adjectival form which should be equivocal', resulting in 'a strain of adjectives in *-ical* which is one of the notes of the literature of the sixteenth and seventeenth centuries, and which has been largely discarded in recent times'. In addition, it has been claimed that the more recent the coinage, the more likely it is to have the *-ic* rather than the *-ical* ending (e.g., Swan 1980:302).

1.1 PRODUCTIVITY OF THE SUFFIXES. To investigate the productivity of *-ic* and *-ical* in different periods, I studied the CD-ROM version of the *OED*, and the first citations of *-ic*/*-ical* adjectives, modelled after Bauer's similar examination of the diachronic productivity of *-ment* (1998:417–19). One can agree with comments on the dangers of inaccuracies in such a research, but it can nevertheless be assumed that the resulting patterns are indicative of the actual productivity (see Görlach 1991:171; Bauer 1998:417). Given the very large number of *-ic*/*-ical* adjectives in the *OED*, it was considered sufficient to examine a small fraction of them. The numbers of first citations of *-ic*/*-ical* adjectives beginning with the randomly chosen letters *a, c, f,* and *u* are presented in Figure 1[1].

The first rather peculiar observation one can make from this figure is that both curves have two peaks, which is also what Bauer (1998:418) discovered with the suffix *-ment*. Bauer concludes from the two-peaked figure of *-ment* formations that 'productivity may change over time in ways that are not directly related to the number of available bases'. The changes in the productivity seem to correspond with the overall numbers of new words coined in these periods: Görlach (1991:137–38) explains that the reasons for the growth period of the English lexicon between 1570 and 1630 were 'the need to express new ideas in English' and 'the subjective desire to enrich the rhe-

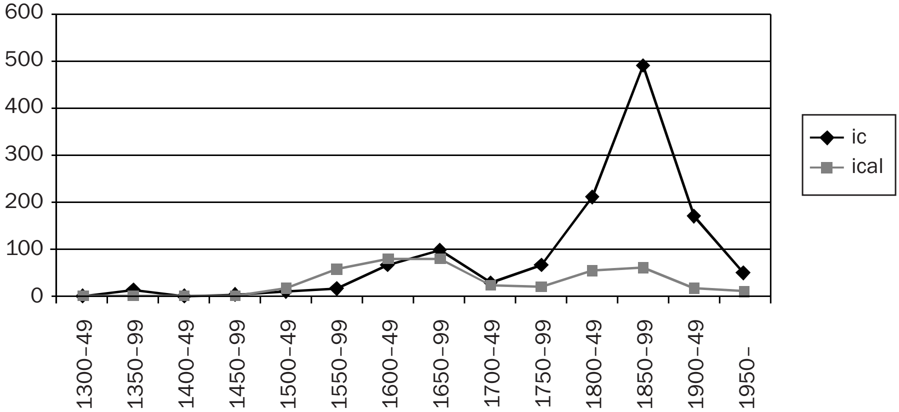


***Figure 1.** The numbers of first citations of words with the suffixes* -ic *and* -ical *in the* OED.

torical potential of the vernacular'. He also remarks that as a result of this period, characterized by linguistic experimentation, there was a great deal of redundancy 'often on the basis of competing derivational patterns'.

It can further be observed from Figure 1 that between 1500 and 1700, roughly the same number of new adjectives were coined with *-ic* and *-ical* endings. Considering especially the latter half of the 16th century, it can be said that the increase in the coinages in *-ical* formations began before that of the shorter forms. During the latter half of the 17th century, new words in *-ic* were coined at an increasingly faster rate, and slightly more new words in *-ic* were recorded than coinages in *-ical*, but the longer suffix still seemed to be productive. In the first half of the 18th century, the numbers of new coinages drop for both suffixes, which again follows the numbers of all new coinages between 1680 and 1780, as reported by Görlach (1991:137). In the 19th century, *-ic* becomes dominant over *-ical* as far as new coinages are concerned, to a great deal because of its high productivity in new scientific terms, especially in chemistry (e.g., *acidic*, *allylic*, *cresylic*, *cyanic*, etc.), with *-ical* forms also resulting in a slight second peak.

1.2. COMPETITION BETWEEN THE SUFFIXES. In the case of competing adjective forms, language has been said to change so as to rid itself of superfluous items, either by discarding one of the forms or by assigning the two forms differing senses (Fowler 1926, s.v. *-ic*, *-ical*; Marchand 1969:241-2). Some comments have been made on the preference for either *-ic* or *-ical* in certain types of adjectives: e.g., according to Sweet (1891:490), *-ic* is preferred in forming adjectives from proper nouns and names of races, as in *Aristophanic*, *Byronic*, and *Slavic*. Marchand (1969:243) singles out a group of adjectives mainly in *-ic* consisting of words with the basic sense 'queer, odd, spleeny', including, e.g., *frenzical*, *nonsensical*, and *quizzical*. In general, whenever there are pairs of adjectives in *-ic/-ical*, it has been claimed that the forms in *-ic* are semantically closer to the base word, whereas with the *-ical* forms the connection is vaguer (Fowler & Fowler 1951, s.v. *-al*), and based on this idea Marchand (1969:242) suggested that scientists are more prone to use the shorter forms than non-scientists (see also Zandvoort 1966:314-5). The first grammarian to comment on the *-ic*/*-ical* variation was Elphinston, who said that '*ic* is a foreign, and *ical* a domestic termination. The former therefore is used upon solemn, the latter on familiar occasions; as *seraph seraphic* or *seraphical*, *microscope microscopic* or *microscopical*. Where the subject then is naturally solemn, the solemn ending prevails; and where familiar, the familiar. So we say almost only *majestic, miltonic*, from *majesty, Milton*, &c. and *whimsical, finical*, from *whimsy, fine*, &c.' (Elphinston 1765:323–24). All these comments are worthy of closer examination, both from a diachronic and synchronic viewpoint. More recently, studies have been also made on semantically differentiated pairs in *-ic*/*-ical*, including Marsden (1985) and Kaunisto (1999).

Görlach comments on the linguistic experimentation of 1570–1630 that 'this proliferation was not cut back until the late seventeenth/eighteenth centuries, as a consequence of natural selection or as a result of grammarians' or lexicographers' prescriptivism' (Görlach 1991:138). As mentioned earlier, in addition to the ending of the 'proliferation', redundant items were dropped from use: Earle (1892:405) observed in his discussion on discarding forms in *-ical* (e.g., *domestical, Germanical, magnifical*) that this practice as regards *-ic*/*-ical* adjectives was not uniform, saying that 'it almost excites a surprise to find that after all we have been rather arbitrary in our discontinuance of some, while we have continued to use others whose case is nowise different', examples of surviving forms being *periodical, statistical, surgical,* etc.

We could then raise the following questions:

1. Is there anything to explain which of the adjective pairs were 'dealt with' by either dropping one of the forms or by semantic differentiation? Why would, for example, semantic differentiation have been a better or more natural way to get rid of redundancy in the case of some adjectives than with some others?
2. If one of the forms is dropped from use, why is it that with some adjectives it is the form in *-ic* (e.g., we do not usually use the form *mathematic*, but *mathematical*), and then with some others the *-ical* form is dropped (e.g., *tragic* instead of *tragical*)?
3. How drastic or quick were the changes in the use of such adjective pairs?

2. THE USE OF THE SUFFIXES IN THE CHADWYCK-HEALEY *LITERATURE ONLINE* CORPUS. To investigate the changes in the use of *-ic*/*-ical* adjective pairs in earlier centuries, the most suitable material to use at the moment is the prose section of the Chadwyck-Healey *Literature Online* text collection, a commercially available corpus. This is simply because there are no other corpora available from this period that would be large enough—the prose section of the *Literature Online* contains altogether approximately 33 million words[2]. What needs to be mentioned is that the choice of this material for the study means, of course, that we are only dealing with fiction, and the results may not be exactly the same if texts of other genres were examined.

	majestic	*majestical*	**Number of authors using both forms**
1550–99	1 (1)	14 (10)	
1600–49	15 (7)	19 (8)	2
1650–99	65 (15)	3 (3)	1
1700–49	33 (9)	—	
1750–99	70 (21)	—	
1800–39	149 (22)	2 (2)	2

Table 1. *The numbers of occurrences of* majestic/majestical *in* Literature Online.

The prose section of the Chadwyck-Healey collection covers the period between 1513 and 1839. As the corpus was obviously not compiled especially to meet the requirements of linguistic corpus research, it may not be ideal as regards its structure. The use of *-ic/-ical* pairs in the corpus was analysed in blocks of 50 years. It is important to note the numbers of books and authors included in the corpus: the numbers of books in different 50 year blocks varied from 43 to 62, and the numbers of authors from 24 to 35 (a curious exception is the period 1700–49, which has 55 books by only 12 authors). In addition, texts by anonymous writers were not included in the analysis. Partly as a result of this decision, the number of books/authors in the corpus from the first half of the 16th century was so small that this period was left out of the analysis entirely.

After making online searches of *-ic/-ical* adjectives, based largely on Lehnert's (1971) reverse dictionary, I studied the use of those adjective pairs which were the most frequent in the corpus, i.e., occurring frequently enough so that observations could be made on them. Thus the analysis did not consider adjectives which did not have an existing counterpart in *-ic* or *-ical*. The analysed adjectives include *angelic(al), authentic(al), comic(al), domestic(al), fantastic(al), heroic(al), magic(al), majestic(al), philosophic(al), poetic(al), tragic(al),* and *tyrannic(al).* Of these adjectives, *authentic/authentical* had the smallest numbers of occurrences, with 80 instances of *authentic* and 27 instances of *authentical*. Pairs with a smaller number of total occurrences in the corpus were excluded from the analysis. It has to be noted again that if other genres were examined, the most frequently occurring *-ic/-ical* adjective pairs could perhaps be different.

Firstly, attention could be paid to the numbers of occurrences of *majestic/majestical* in the *Literature Online* corpus, presented in Table 1. In Table 1, the numbers given under the adjective forms denote the absolute number of occurrences of the forms in different fifty-year periods, and the numbers in parentheses denote the numbers of authors using the adjectives. On the right hand side of the table are the numbers of those authors who were using both forms. As far as the semantics of the adjectives here are concerned, both *majestic* and *majestical* were used in the same meanings, as can be seen in examples (1) a and b (see also *OED*, s.vv.), found in the *Literature Online* corpus:

	authentic	*authentical*	Number of authors using both forms
1550–99	—	11 (6)	
1600–49	3 (2)	10 (6)	
1650–99	24 (11)	6 (3)	
1700–49	16 (5)	—	
1750–99	15 (11)	—	
1800–39	22 (9)	—	

Table 2. *The numbers of occurrences of* authentic/authentical *in* Literature Online.

(1) a. Though her Beauty had (while present) suspended Cha-abas Passion towards the Sultana for that time, he had no sooner retired from her Company, but that the Sultana presented her self, to his Imagination, with all those Perfections, and that Majestick Grace which always accompany her. (Peter Bellon, *The Court Secret* (1689), 332)

b. …they arriued shortly at the place of triumph, into which entring, bowing her selfe vnto her horse necke, vnto those she first met, she passed her horse by ye tilt, with such a countenance, and maiestical grace, that euery one admired, both her persone, and good demeanour… (Henry Roberts, *Honovrs Conquest* (1598), 137)

In the periods when both forms were actively used, there were no signs in the texts of clear preference for the use of either form in any particular sense or collocation. Similarly, there seemed to be no significant difference as to the positioning of the adjectives: both forms appeared roughly with the same relation in attributive and predicative positions. It can be seen from Table 1 that in the latter half of the 16th century, the *-ical* form was more common—and this seems to be the case with most of the *-ic/-ical* adjective pairs examined—but later in the 17th century, the use of the shorter form became more popular, and the *-ical* forms less frequent towards the end of the 17th century, finally almost disappearing. In present-day English, *majestical* is indeed quite rare. So here we would have evidence of a clear shift from *-ical* to *-ic* forms, with no apparent trends of semantic differentiation, but merely one of dropping the forms in *-ical* from use.

Interestingly enough, the same kind of change—both as regards the synonymous use of the forms and the time when one can perceive a shift in favor of the shorter forms—can be seen, for example, with the adjective pairs *authentic/authentical* and *heroic/heroical,* as is evident from Tables 2 and 3.

A notable feature in Table 3 is the numbers of authors using both forms, especially in the period 1600-49, when almost all of the writers who used the adjectives *heroic/heroical* were using them interchangeably, as e.g., Alexander Hart in examples (2)a and b:

	heroic	*heroical*	Number of authors using both forms
1550–99	5 (2)	48 (15)	
1600–49	26 (8)	36 (7)	6
1650–99	79 (17)	40 (7)	6
1700–49	69 (9)	3 (1)	1
1750–99	141 (21)	3 (3)	3
1800–39	145 (20)	3 (3)	3

Table 3. *The numbers of occurrences of* heroic/heroical *in* Literature Online.

	tragic	*tragical*	Number of authors using both forms
1550–99	16 (7)	69 (19)	5
1600–49	17 (7)	61 (11)	5
1650–99	12 (7)	51 (18)	3
1700–49	20 (4)	34 (8)	4
1750–99	38 (10)	53 (16)	5
1800–39	79 (11)	27 (13)	7

Table 4. *The numbers of occurrences of* tragic/tragical *in* Literature Online.

(2) a. This grave Senator had a Sonne who bearing his name, disagreed not from his properties, but punctually had his Fathers heroicke inclinations, which made him as happy as meritorious, and was no lesse honoured of all men… (Alexander Hart, *Alexto and Angelina* (1640), 19–20)

b. …Sandrico for the better satisfaction of his friend and himselfe, begun to enquire and demand of the Inhabitants what they could informe him concerning the heroicall inclination of this Roman Goddesse,… (Alexander Hart, *Alexto and Angelina* (1640), 56)

It seems justifiable to say then that with some adjective pairs ending in *-ic* and *-ical*, there was a noticeable shift from the use of *-ical* forms to the *-ic* forms in the 17th century, although during this period, as could be concluded from the *OED* chart in Figure 1, new words in *-ical* were still coined. This is particularly interesting in the light of the comments made by Earle (1892), as the tendency to discard the existing forms in *-ical* appears to have started earlier than he had assumed. In fact, it appears to have begun before the use of the suffix as a productive derivational pattern had gone out of fashion.

On the other hand, we do not get similar results to these when examining pairs such as *tragic/tragical* and *tyrannic/tyrannical*, as seen in Table 4 and Table 5 (overleaf), conforming with Earle's observation of the adjective pairs' behaviour not being uniform.

	tyrannic	*tyrannical*	Number of authors using both forms
1550–99	—	28 (8)	
1600–49	—	2 (2)	
1650–99	7 (4)	46 (13)	1
1700–49	8 (3)	19 (6)	2
1750–99	9 (8)	44 (20)	7
1800–39	12 (5)	44 (12)	4

Table 5. *The numbers of occurrences of* tyrannic/tyrannical *in* Literature Online.

	philosophic	*philosophical*	Number of authors using both forms
1550–99	—	12 (5)	
1600–49	—	6 (3)	6
1650–99	1 (1)	23 (8)	6
1700–49	4 (4)	21 (7)	1
1750–99	54 (19)	39 (18)	3
1800–39	45 (9)	78 (17)	3

Table 6. *The numbers of occurrences of* philosophic/philosophical *in* Literature Online.

In the case of *tragic/tragical*, the longer form is in present-day English rare when compared to *tragic*. From a temporal viewpoint, however, the developments with this word pair do not seem to go along the same lines as with the adjective pairs discussed earlier, as the form in *-ical* appears to have survived well into the 19th century. With *tyrannic/tyrannical*, the *-ic* form enters the scene in the latter part of the 17th century, but the status of the longer form never became threatened, and today *tyrannical* is clearly the favored form.

Even more interesting changes can be found in the use of *philosophic/philosophical*, the numbers occurrences given in Table 6. In the latter half of the 18th century, the shorter form became more common, but during this period the two forms were used differently: *philosophic* in this period was more often used in a wider, more popularized meaning, appearing in phrases such as *a philosophic manner* or *a philosophic look/expression*. When referring to things pertaining to philosophy as a science, with collocating words such as *research*, *theories* or *argumentation*, *philosophical* was clearly the preferred form. After that period, however, this semantic distinction began to disappear, and *philosophical* became the more common form in all senses used—and in present-day English, one can still occasionally come across instances of the shorter form, but *philosophical* is considerably more frequent. In this case we would then actually encounter both processes of getting rid of redundancy, semantic differentiation and dropping one of the forms from use.

One may wonder why adjective pairs seem to differ so much in their behavior as regards either differentiation between the forms or dropping one of the forms. If there was a shift from *-ical* to *-ic* in the 17th century, why did it affect only some adjective pairs, and not all of them? One possibility could be in the meanings of the base words, as proposed by Elphinston (1765). While nothing in the use of the examined adjectives in the corpus suggests that the use of either the *-ic* or *-ical* form depended in any way on the 'familiarity' or 'solemn nature' of the occasion (as Elphinston also argued), the semantic aspects of the words themselves do provide a common feature. For example, *angelic(al), heroic(al),* and *majestic(al)* seem to have more in common with each other than with *philosophic(al)* or *poetic(al),* the shared component being the semantic aspect of nobility. If this indeed were the case, an interesting conclusion would be that the meaning, the solemn nature of the base word itself might have had an effect in the choice of how to get rid of redundancy: not semantic differentiation, but dropping the forms in *-ical* from use. Although Elphinston did not specifically comment on the language actually changing as regards the use of the two suffixes, his observations on the matter—somewhat surprisingly not referred to since by lexicographers or grammarians—would seem to find partial support from the material examined.

Based on the *Literature Online* material alone, it might be early to suggest classifications of *-ic*/*-ical* adjectives according to the changes in their use, but presented below is a preliminary suggestion of such a classification, showing that adjectives having the semantic component of nobility could, in fact, be observed as behaving in the same way. This tendency is further strengthened in the corpus by the adjective pairs *prophetic(al)* and *seraphic(al),* which, although not as frequent as the other pairs in group 1, nevertheless show a similar shift from *-ical* to *-ic*.

I. shift from *-ical* to *-ic* in the 17th century or early 18th century—
   NO SEMANTIC DIFFERENTIATION
   a. base word has the semantic characteristic of nobility or grace: *angelic(al), authentic(al), heroic(al), majestic(al), prophetic(al), seraphic(al)*
   b. others: *domestic(al)*

II. # of occurrences in *-ic* increases (in the 18th century)—
   SEMANTIC DIFFERENTIATION—both forms remain in use
   a. the form in *-ic* stronger/usual: *comic(al), fantastic(al), magic(al)*
   b. the form in *-ical* stronger/usual: *poetic(al)*

III. # of occurrences in *-ic* increases, but the *-ical* form remains the usual form–
   NO SEMANTIC DIFFERENTIATION
   tyrannic(al)

IV. the form in *-ic* becomes more frequent—SEMANTIC DIFFERENTIATION—
   followed by a shift back to *-ical*, with NO SEMANTIC DIFFERENTIATION
   philosophic(al)

V. shift from *-ical* to *-ic* after the first half of the 19th century—
   NO SEMANTIC DIFFERENTIATION
   tragic(al)

3. CONCLUSION. It could be said that obviously more research has to be done on the *-ic/-ical* adjectives in the Early Modern English period as well as other periods so as to include more adjective pairs in the classification—and there would probably be even more classes than the ones suggested here. The findings from the *Literature Online* corpus does not obviously give us answers for all of the questions about the use of *-ic/-ical* adjectives, but they do suggest that there have been a number of different currents of change. The observation made on adjectives such as *angelic(al), heroic(al)*, and *majestic(al)*, sheds new light on the matter of competition between these two derivational patterns in that in cases where both forms have already existed, the semantic aspects of the adjectives themselves may be considered to have affected the application of one of language's 'self-purifying' tendencies, in this case attrition instead of differentiation.

---

[1] Excluded from the investigation were of course those words which actually were prefixed forms of already existing words in *-ic/-ical*.

[2] This figure refers to the prose fiction databases of the *Literature Online* as they were in February 2000—the 19th century fiction database has been since completed to include works from 1840 to 1900.

REFERENCES

BAUER, LAURIE. 1998. Is there a class of neoclassical compounds, and if so is it productive? *Linguistics* 3:403–22.

BELLON, PETER. 1689. *The court secret*. London: R.E. for R. Baldwin. [From the electronic version in *Chadwyck-Healey literature online*.]

BIBER, DOUGLAS, SUSAN CONRAD & RANDI REPPEN. 1998. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press.

Chadwyck-Healey literature online. 1996-2000. Cambridge: Chadwyck-Healey Ltd (Bell & Howell Information and Learning Company).

EARLE, JOHN. 1892. *The philosophy of the English tongue,* 5th edition. Oxford: Clarendon Press.

ELPHINSTON, JAMES. 1765. *The principles of the English language digested*. London: James Bettenham. (out of print—available in the British Library, London).

FOWLER, H.W. 1926. *Modern English usage*. Oxford: Clarendon Press.

——— & F.G. FOWLER, eds. 1951. *The Concise Oxford dictionary of current English,* 4th edition, revised by E. McIntosh. Oxford: Clarendon Press.

GÖRLACH, MANFRED. 1991. *Introduction to early modern English.* Cambridge: Cambridge University Press.

HART, ALEXANDER. 1640. *Alexto and Angelina*. London: B.A. and T.F. for Nich. [From the electronic version in *Chadwyck-Healey literature online.*]

KAUNISTO, MARK. 1999. *Electric/electrical* and *classic/classical*: Variation between the suffixes *-ic* and *-ical*. *English studies* 4:343–70.

KENNEDY, GRAEME. 1998. *An introduction to corpus linguistics.* London: Longman.

LEHNERT, MARTIN. 1971. *Reverse dictionary of present-day English*. Leipzig: VEB Verlag Enzyklopädie Leipzig.

MARCHAND, HANS. 1969. *The categories and types of present-day English word-formation,* 2nd edition. Munich: C.H. Beck.

MARSDEN, P.H. 1985. Adjective pairs in *-ic* and *-ical*. *Lebende Sprachen* 1:26–33.

MURRAY, J.A.H., H. BRADLEY, W.A. CRAIGIE *&* G.T. ONIONS, eds. 1992. *The Oxford English dictionary*, second edition on compact disc, prepared by I.A. Simpson and E.S.C. Weiner. Oxford: Oxford University Press and Software. (*OED*)

ROBERTS, HENRY. 1598. *Honovrs conqueſt*. London: Thomas Creede. [From the electronic version in *Chadwyck-Healey literature online*.]

SWAN, MICHAEL. 1980. *Practical English usage*. Oxford: Oxford University Press.

SWEET, HENRY. 1891. *A new English grammar. Part I: Introduction, phonology, and accidence.* Oxford: Clarendon Press.

ZANDVOORT, R.W. 1966. *A handbook of English grammar,* 4th edition. London: Longmans.

~

FRENCH *ÇA* AND THE DYNAMICS OF REFERENCE

Michel Achard
Rice University

introduction. The treatment of demonstrative pronouns in French is usually simple and uncontroversial. This statement by Grevisse (1986: 1054) is representative of the analyses found in traditional grammars, as well as early linguistic accounts:

> ‘*Les pronoms demonstratifs désignent un être ou une chose en les situant dans l'espace, éventuellement avec un geste à l'appui (fonction* déictique*). Ils peuvent aussi renvoyer à un terme qui précède (fonction* anaphorique*) ou qui suit (fonction* cataphorique*) dans le contexte.*’

> ‘Demonstrative pronouns designate a person or a thing by situating them in space, sometimes with an accompanying gesture (*deictic* function); they can also refer to a preceding term (*anaphoric* function) or a following term (*cataphoric* function) in the context’.

In this paper, I argue that the French demonstrative pronoun *ça* (which will be alternatively translated by ‘this’, ‘that’, or ‘it’) does not merely point to an established referent, but creates that referent from the surrounding discourse. In this view, the pronoun's presence instructs the hearer to organize the context of the utterance in a highly specific fashion. In order to argue this position, I examined 464 instances of *ça* from a corpus of newspaper (*Le Monde*) articles published between 1988 and 1990. The analysis proposed here is not intended to be exhaustive of all the possible senses of the pronoun, but merely indicative of a methodology to approach its investigation.

1. a first attempt at the meaning of *ça*. A strictly referential analysis of *ça* would consider the pronoun's discourse function as a direct extension of its deictic value. It is untenable because in numerous instances, *ça* doesn't anaphorically refer to an existing entity. The first argument against a referential account is provided by the contrasts that exist between true anaphors (personal pronouns) and *ça* in related situations. Contrasts such as the ones illustrated in (1) and (2) are well studied in the literature, most notably by Cadiot (1988) and Carlier (1996); the example in (1) is adapted from Carlier 1996:

(1) a. ***Les gosses, ils** se lèvent tôt*
The kids, they get up early

b. ***Les gosses, ça** se lève tôt*
Kids, they get up early

(2) a. ***Les Français, ils** se plaignent tout le temps*
The French, they are always complaining
b. ***Les Français, ça** se plaint tout le temps*
French people, they complain all the time

In (1)a and (2)a, the personal pronoun *ils* 'they' is strictly anaphoric of its referent because i) the agreement on the verb matches up with the gender and number of the referent, and ii) the preferred reading of the utterance is specific. For example, in the preferred reading of (1)a, the speaker has specific kids in mind, and it is about those kids that the predication of *se lever tôt* 'get up early' is made. The examples in (1)b and (2)b illustrate what Cadiot (1988) calls *décalages referentiels* 'referential shifts'. The subject/verb agreement does not match up with the referent's number (it is in the singular), and the reading of the utterance is necessarily generic. This generic reading is confirmed in (3) by the restrictions that exist on the determiner allowed in the NP that denotes *ça*'s referent. (Carlier 1996).

(3) a. ***Le/mon/les gosses du voisin, ils** se lèvent tôt*
The/my/the neighbor's kids, they get up early
b. ****Le/mon/les gosses du voisin, ça** se lève tôt le matin*
The/my/the neighbor's kids, that gets up early

The second argument comes from the cases where the pronoun's referent is not immediately identifiable, as illustrated in (4):

(4) *J'avais gardé de bons copains du temps de l'Opéra, dont un qui était passé chez Cuevas. Il m'a attrappé dans un bar de la rive gauche et m'a conseillé la mode, m'a envoyé présenter mes dessins. A l'époque, je m'habillais beaucoup, j'étais presque un personnage avec des avions, des nuages d'or et d'argent découpés sur le dos de mes blousons. Je ne savais rien, j'ai préparé un dossier,* ***ça*** *a marché. J'ai appris comme* ***ça****, et je n'ai pas arrêté.*
'I had kept good friends from the Opéra days, including one who had gone to Cuevas. He caught me in a bar on the left bank and advised me to get into fashion, and sent me to show my drawings. At the time, I would dress up a lot, I was almost a character with planes, gold and silver clouds cut outs in the back of my jackets. I didn't know anything, I prepared a portfolio, **it** worked out. I learned like **that**, and I never stopped.'

Intuitively, the referent of the first *ça* cannot be the nominal *dossier* 'portfolio', because a portfolio cannot be said to be working or not. The pronoun therefore doesn't refer to a nominal directly available in the current discourse. If *ça* cannot be considered a direct anaphor to its referent, how can we characterize its meaning? Or, to rephrase the question in Cognitive Grammar terms (Langacker 1987a, 1991)[1], what does *ça* profile? I would like to suggest that *ça* provides the hearer with the instructions to create a specific region within the space provided by the context of the utterance. To put it differently, within the space of the context of the utterance, *ça*'s presence instructs the hearer to construe a group of entities as interconnected (Langacker 1987a, 1987b).[1] The entities to be considered together are not always made specific, and the pronoun's presence often requires the hearer to actively manipulate the current discourse in order to achieve the appropriate construal. In that view, *ça*'s presence enables the speaker and hearer to interactively sequence and group certain elements of the text for communicative purposes.

2. *ÇA* AND THE NOTION OF CREATED REGION. This hypothesis accounts at the same time for the pronoun's deictic meaning as well as its discourse sense.

2.1. DEICTIC SENSE. Compare the example in (5)a to the one in (5)b. The context in (5) is that of a child expressing her preference in front of a pastry shop window displaying cakes.

(5) a. *Je veux ça* b. *Je veux celui là*
I want that I want this one

In both cases, the child needs to point toward the desired entity to get her point across, but assuming that her gesture is accurate, her utterance will produce the same result in both instances. However, the pronoun's referent is accessed differently in the two situations. In (5)b, the type *gâteau* 'cake' has either already been introduced, or it is implicit from the discourse situation, but it is shared between the child and her interlocutor. The child is pointing to a particular instance of that established type as the object of her desires. In (5)a, there is no specific type shared between the interlocutors. The child is pointing to the cake as some portion of an undifferentiated mass. She is not accessing that mass as an example of any category, but in a sort of pre-categorized way, before the separation in the different types is made. *Ça* therefore profiles a region of undistinguished (pre-categorized) physical matter, which represents the definition of an object before it is characterized as a specific type. The specific nature of the object is recoverable by the immediate context, hence the importance of the accompanying gesture.

2.2. REFERENTIAL SHIFTS. Consider the referential shift construction illustrated in (6):

(6) ***La voiture, ça*** *me plait*
Cars, I like them

We have already seen that the nominal *la voiture* cannot be considered the referent of *ça* per se. It is also obvious that *ça*'s referent is closely related to that nominal. Consistent with the hypothesis proposed in the previous section, I claim that the pronoun's referent is a more abstract nominal composed of a specific region created around *la voiture*. In (6), the predicate *plaire* profiles the conceptualizer's reaction to several possible kinds of activities associated with cars. It therefore seems reasonable to say that the mention of the nominal *voiture* evokes the set of activities associated with it. Activity is understood in its broadest sense that includes events (and possibly full propositions) where the nominal is a participant. *Ça* profiles the abstract region defined by construing these activities as a set of interconnected entities.

Note that in these examples, there is no restriction on the possible nature of the extracted nominal. Definites and proper nouns are perfectly felicitous, as shown in (7) and (8).

(7) *Mon frère, **ça** m'intéresse*
My brother, that interests me

(8) *Tenez, sans aller si loin, il était pas bien mon Jacquot à l'Hotel de Ville ? Non, **Matignon, ça** la foutait mieux, **c'**était plus classe question adresse.*
There we go, no need to look further. City Hall wasn't good enough for Jack baby? No, Matignon, **that** sounded better, as an address, **that** had more class.

The example in (7) can only be fully understood in context. *Ça* profiles the abstract region of activities, behaviors, beliefs etc. that surround the speaker's brother, in relation to which the predicate can be asserted. In (8), the proper name *Matignon* refers to the French prime minister's residence, and thus metonymically stands for his function, with all the considerations of money, power, prestige that it conveys. *Ça*'s presence provides directions to construe these considerations as a set of interconnected entities (i.e. an abstract region).

Let us now turn to the constructions briefly introduced in (1) and (2). Additional examples are presented in (9) and (10).

(9) *Ce que vous pouvez être susceptibles, les enfants, c'est pas croyable. Tant que j'égratigne le voisin, ça va, vous m'envoyez des gerbes de fleurs: Ce que t'es rigolote ! T'es vraiment la reine. Génial, ton billet, l'autre jour, sur les paysans. On l'a lu en famille, on se marrait comme des baleines. Mais quand c'est vous que je taquine, ça va pas, je reçois des giclées de bile : Ce que t'es conne ! T'es vraiment la reine. Nulle, ta crotte, l'autre jour, sur les agriculteurs. Un copain me l'a lue au téléphone, j'en croyais pas mes oreilles. Les pires, c'est les Corses, les juifs, les Arabes, les chasseurs et les enseignants. C'est ombrageux, **l'enseignant**, on n'a pas idée. **Ça** sourcille, **ça** se hérisse, **ça** se vexe pour un rien. Là, en ce moment, je **les** ai tous sur le paletot.*
'How succeptible can you be, guys. I can't believe it. As long as I scratch your neighbor, everything is fine, you send me flowers: How funny you are! You are really the queen. Superb, your note the other day about peasants. The whole family read it and we couldn't stop cracking up. But when it's you I tease, we have a problem. You cover me with bile: What an idiot you are! You are really the queen. Zero, your piece of trash about farmers the other day. A friend read it to me over the

phone, I couldn't believe my ears. The worst are the Corsicans, the Jews, the Arabs, the hunters and the teachers. **That** broods, a teacher, one wouldn't believe. **That** frowns, **that** gets all prickly, **that** gets offended for nothing. Right now, I have them all against me.'

(10) ***Etre ministre aujourd'hui*** *c'est d'abord, semble-t-il, apprendre à faire taire l'homme de gauche qui sommeille en soi.* ***Un ministre*****,** ***ça*** *ouvre sa gueule à la rigueur (drôle d'expression) pour parler de " justice sociale ", de " partage ", de " solidarité " et de " pacte de croissance ", mais si* ***ça*** *ne veut pas démissionner,* ***ça*** *doit la boucler devant des salariés agitant leur fiche de paie sur laquelle les grands mots lancés depuis huit ans n'ont pas fait fleurir beaucoup de zéros.*
'To be a secretary of state nowadays is primarily, it seems, to learn to gag the left wing person inside of us. A secretary of state, **that** opens its mouth if necessary to talk about "social justice" , about "sharing", about "solidarity", and "growth plan", but if **that** doesn't want to resign, **that** must shut up in front of workers who wave their pay stubs on which the rhetoric of the past eight years hasn't added many zeros.'

The examples in (9) and (10) are similar to the ones presented in (6)–(8) with the exception that the predications made about the relevant nominals represent definitional properties or attributes of those nominals. For example, the category *enseignant* in (9) includes the behavioral traits described in the following predications. I propose that in these constructions, *ça* profiles a specific region created around the evoked category. This region is composed of the sum of the elements that provide the conceptualizer with her knowledge of the category, that is to say the facts, opinions, or other beliefs that arise in the conceptualizer when it is evoked. This knowledge of the category enables the speaker to pronounce the judgement expressed in the following predicates. To continue with (9), the category *enseignants* evokes the facts, beliefs, or even sensations or memories about its members that constitute the speaker's knowledge of the category. This knowledge is expressed by the predicates *être ombrageux* 'to brood', *sourciller* to frown', *se hérisser* 'to get prickly', *se vexer pour un rien* 'to get offended for nothing'.[2]

2.3. *ÇA* AND REALITY. In (11), the political figure *Alain Poher* was famous for maintaining a high profile in the senate for a very long time.

(11) *Poher,* ***ça*** *suffit ! Comment ils osent ?* ***Ça*** *fait jamais que vingt ans qu'il y trône, au palais du Luxembourg, le pauvre chéri.*
'Poher, **that**'s enough! How dare they? **It**'s only been twenty years he's been reigning at the Luxemburg palace the poor dear.'

The exact nature of the abstract region that surrounds the nominal *Poher* is difficult to characterize precisely. One possible solution would be to define it as the sum of the senator's accomplishments. However, it doesn't seem convincing in this context, because political successes and failures do not constitute the topic of discussion. Rather, what is being commented on is the amount of time during which Poher has been part of the political scene. Because the passing of time is so crucial to the understanding of the utterance, I suggest that the abstract region anchored around the extraposed nominal represents the specific portion of evolutionary momentum (Langacker 1991) that prominently contains Poher. The second occurrence of *ça* provides additional details by indicating the exact duration of that sub-part of reality.[3]

Instead of a person, the section of reality *ça* evokes can be anchored around a specific situation as in (12).

(12) *Et puis, un soir de janvier 1989, le téléphone a sonné. C'est Daniel qui a décroché, là-bas près du canapé rose. " Manuela, quoi Manuela ? - Elle vivait avec un jeune homme.* ***Ça*** *s'est mal passé entre eux. Il l'a tuée. Venez vite. On vous attend. "*
And then, one evening of January 1989, the telephone rang. Daniel picked up, over there by the pink sofa. 'Manuela, what about her?—She was living with a young man. **It** went sour between them. He killed her. Come quickly. We are waiting for you.'

In (12), *ça* evokes the sub-part of reality that encompasses the common life of Manuela and her boyfriend. Its contents is available to the hearer as part of her world knowledge of what the life of a young couple might entail. *Ça* instructs the hearer to connect together the different facets of what she knows about life as a couple to create a nominal construal of the young lovers' experience. The ensuing event or proposition can then be predicated of that created nominal.

In other cases, *ça* evokes the sub-part of reality that includes a particular event or proposition, as illustrated in (13) and (14).

(13) *Nous vendons en moyenne quinze mille exemplaires d'un disque.* Passion, *d'Arvo Pärt, en est à trente mille.* ***Ça*** *s'est fait de façon non préméditée. Je ne réfléchis jamais en terme de rendement ou de budget.*
On average, we sell fifteen thousand copies of an album. *Passion* by Arvo Pärt has sold thirty thousand so far. **It** happened unexpectedly. I never think in terms of revenue or budget.

(14) *Il faut dire que le maire de Beaucaire est lui-même instituteur depuis vingt-huit ans, mari et père d'institutrices. C'est en partie pour des raisons "éducatives" qu'il justifie les mesures prises à l'encontre d'enfants immigrés. "J'ai dans ma classe quatre enfants maghrébins sur vingt-huit.* ***Ça*** *ne pose aucun problème. Mais, quand on arrive, comme c'est bien souvent le cas à Beaucaire, à plus de 50 %,* ***ça*** *devient impossible".*
One must say that the mayor of Beaucaire has been a school teacher himself for twenty eight years. He is also the husband and father of school teachers. It is partly

for 'educational' reasons that he justifies the measures taken against immigrant children. 'In my class, I have 4 children from the Maghreb out of 28 students. **That** doesn't pose any problems. But when you reach a total of over 50% as it is often the case in Beaucaire, **that** becomes impossible.'

In (13), *ça* represents the segment of evolutionary momentum that brought about the proposition *Passion, d'Arvo Pärt, en est à trente mille* '*Passion* by Arvo Pärt has sold thirty thousand (copies) so far' as part of present reality. In (14), the communicative purpose of the paragraph is to contrast two kinds of realities differentiated by the propositions they contain. The first *ça* evokes the sub-part of reality that includes the proposition *J'ai dans ma classe 4 enfants Maghrébins sur 28* 'In my class, I have 4 children from the Maghreb out of 28 students'. The second *ça* evokes the sub-part of reality that includes a different proposition, namely *quand on arrive à plus de 50%* 'when you reach a total of over 50%'.

Finally, the scope of the sub-part of reality *ça* evokes can be wide enough to coincide with present reality, as shown in (15)–(17).

(15) *Une femme est debout dans l'entrée, blonde, les cheveux sur les épaules. Une apparition. "Tu me reconnais ? ", demande-t-elle. "Oui, répond Suzy. Mais comment ça se fait que t'es là ?"*
A woman is standing in the hallway, blond, her hair shoulder length. An apparition. 'Do you recognize me?' she asks. 'Yes, Suzy answers. But how come you are here?'

(16) *Dans dix jours, Mr Gorbatchev présente devant le Comité central un plan de démocratisation accélérée du parti et du pays. Plus* ***ça*** *va mal, plus il veut aller de l'avant, convaincu qu'il est maintenant que seule la radicalisation de la perestroika peut empêcher son échec, puisqu'il ne faut pas freiner mais accélérer dans les virages.*
In ten days, M. Gorbachev will present a plan of accelerated democratization for the party and the country in front of the Central Committee. The worst **it** goes, the more he wants to go forward, convinced as he is that only the radicalization of Perestroika can prevent its failure, since one should not break but accelerate in the middle of a turn.

(17) *Olivier a donc été appelé à la barre. Père d'un enfant depuis un mois, c'est pourtant encore un adolescent qui a eu dix-neuf ans en septembre dernier. Il est si intimidé qu'il murmure ses réponses au président Jean-François Périé: " J'avais des problèmes, je préférais mourir… ", balbutie le jeune homme. " Et maintenant,* ***ça*** *va ? " s'inquiète le magistrat. Après un silence, Olivier soupire : "* ***Ça*** *va... "*
So Olivier was called to testify. The father of a child for a month, he is, however still a teenager who turned nineteen last September. He is so intimidated that he mumbles his answers to president Jean François Périé: 'I had problems, I preferred

to die' the young man whispers. 'And now, is **it** OK?' the judge asks. After a pause, Olivier sighs '**It**'s OK...'

In these three examples, *ça* evokes present reality, that is to say the sum of events and propositions that combined to produce the state of the world we are currently witnessing. For example in (15), Suzy is expressing surprise at the current state of events that include the presence of a person she didn't anticipate to see.

Specific aspects of reality can be made particularly salient in their construal by *ça*. For example, the pronoun's referent frequently profiles the precise length of time reality has included a certain event or proposition. Such cases are presented in (18)-(20).

(18) *Mon Jules, quel Jules ? **Ça** va faire deux ans qu'on est plus ensemble.*
My man, what man? **It** will be two years we haven't been together

(19) *Et puis, dès qu'il a appris la nomination d'Antoine Vitez à la tête de la Comédie-Française, ce jeune comédien free lance s'est dit qu'il était peut-être temps de s'engager plus que pour un rôle : " **Ça** fait huit ans que je fais ce métier et je ne m'étais jamais engagé à long terme. C'est pour ça que je n'ai pas fait le Conservatoire ou d'école.*
And then, as soon as he learned Antoine Vitez's was the new director of the Comédie Française, this young free lance comedian said to himself that maybe it was time to commit himself for more than one role: '**It**'s been more than eight years that I've been in this business and I had never committed myself for the long run. That's why I never went to the Conservatory, or any other school.'

(20) *" Alors, **ça** y est, Edu, tu te remets à chanter ? " C'est avec l'humour des gens de Rio, toujours prompts à raconter les pires histoires en se moquant d'eux-mêmes, comme s'ils étaient les acteurs d'une comédie de boulevard, qu'Edu Lobo évoque ses vingt-cinq ans de carrière : celle d'un compositeur qui voudrait bien n'avoir qu'à composer, mais est contraint par la vie, les autres, ou lui-même, à remonter sur scène.*
'So, that's it Edu, you are singing again?' It is with the humor of the people of Rio, always ready to tell the worst stories and make fun of themselves, as if they were the actors in a Vaudeville, that Edu Lobo evokes his twenty five year career: the career of a composer who wished he could have remained only a composer, but who was forced by life, other people or himself to start performing again.

2.4. COGNITIVE MANIPULATIONS. So far, the analysis has concentrated on the nature of the region *ça* creates out of the immediate context. This section examines the way in which *ça*'s presence signals the cognitive manipulation (summarization, abstraction, inferrencing) of information already available in the context or in the hearer's world knowledge. In (21) and (22), the pronoun's presence instructs the hearer to summarize the information at her disposal.

(21) *Les affaires de Manuela ? Elles étaient sous scellés, intouchables. Les seules parcelles de réalité, les seules preuves symboliques de la mort de Manuela, Maité Gourjault les a puisées alors dans ces rares images : un policier dans un souterrain, un avocat à qui elle donne un chèque, deux rendez-vous chez le juge d'instruction, l'incinération d'une caisse de bois anonyme et quelques papiers qu'elle a pu feuilleter. Et puis il y eut les lettres du jeune homme écrites depuis sa prison. " Bonjour Maité, dis-moi comment tu vois* ***tout ça vu de l'extérieur****. Pour ma part, tout reste inexplicable."*
Manuela's stuff? It was sealed, untouchable. The only fragments of reality, the only symbolic proofs of Manuela's death, Maité Gourjault dug them out of these rare images: a policeman in a tunnel, a lawyer to whom she gave a check, two appointments with the judge, the cremation of an anonymous wooden box and a few papers she could flip through. And then there are the young man's letters from his jail cell. 'Hello Maité tell me how you see all **that** from the outside. For my part, everything remains inexplicable'.

(22) *On plonge dans Yellow Submarine avec une facilité qui, au début, déconcerte. Le délire visuel coordonné par George Dunning reprend tous les clichés de la période de l'op'art, qui fait mal à la tête, au style nouille passé à la peinture orange. Mais, sous la surcharge, le charme agit parce que, en accumulant une douzaine de trouvailles à chaque plan, on finit par tomber juste assez souvent. Et, de collages en mises en abime, de labyrinthes en solarisation, on finit par retrouver la musique, univers poétique du groupe.* ***Tout ça*** *grâce à une bonne volonté sans arrière-pensée qui fait de cette oeuvre de commande un instantané parfait de l'apogée des Beatles.*
One dives into Yellow Submarine with such ease that it is disconcerting at the beginning. The visual delirium coordinated by George Dunning makes use of all the clichés from the op'art period that gives you a headache to the noodles-covered-in-orange-paint style. But underneath the overload, the charm is effective because when one accumulates a dozen findings per plan, one ends up being right quite often. And from collage to zoom, from labyrinth to solarization, one ends up finding the music, the poetic universe of the group. All **that** thanks to a good will without second thoughts that makes that work on command a perfect snapshot of the peak years of the Beatles.

The example in (21) comes from the same article as the one in (12). *Manuela*, *Maité*'s daughter has been killed by her boyfriend (the *jeune homme*). It is quite difficult to precisely identify the entity that *ça* evokes, but it clearly involves the summary of the circumstances leading to (and including) the current situation. It is interesting to note that this entity is not shared by the participants in the discourse. The circumstances that surround Manuela's death are necessarily different (at least in part) for her murderer and her mother. This is, however, unproblematic for the analysis developed here where *ça* provides the hearer with the instructions to create a summary of the circumstances available to her, whatever those might be. The abstract region profiled by *ça* is composed of the mother's views, beliefs, feelings, and opinions about her daughter's death connected together. In (22), the information to be summarized is specifically provided (at least partly) in the immediate context. *Ça* profiles the summary of the experiences accumulated throughout the show.

The presence of *ça* often involves other conceptual manipulations in addition to summarizing. Some of them are illustrated in (23) and (24).

(23) *Vous avez vu Radio days, de Woody Allen, c'était exactement* ***cela*** *! Je ne peux pas dire mieux ! Cette maison était la mienne, la femme qui ne se marie pas, c'était ma tante !* ***La rue qui donne sur la mer, les voisins, l'école, les parents, le football, le base-ball****, tout* ***ça*** *était ma vie !* "
You saw *Radio days* by Woody Allen, it was exactly **that**! I can't say it any better! That house was mine, the woman who doesn't marry, **it** was my aunt! The street that overlooks the sea, the neighbors, the parents, football, baseball, all **that** was my life!

(24) *Un an et demi après nous devons constater que pour nous, la vie quotidienne ne change pas. En banlieue de Nouméa, à Saint-Louis par exemple,* ***six cents familles wallisiennes vivent encore sans eau courante et sans électricité*** *et, il y a quelques semaines, plus près de Nouméa, dans la vallée du Tir, nous avons dû manifester parce qu'****on avait coupé l'eau à tout un quartier pour sanctionner sept mauvais payeurs****. C'est tous les jours comme* ***ça*** *et, si l'on n'y remédie pas, cela signifie que les accords de Matignon ne sont qu'une mascarade.*
One and a half years later, we must say that for us daily life doesn't change. In the suburbs of Nouméa, in St Louis for example, six hundred Wallisian families still live without running water or electricity, and a few weeks ago, closer to Nouméa, in the Tir valley, we had to demonstrate because the water had been turned off for a whole neighborhood to penalize seven people who hadn't paid their bills. Everyday it's like **that** and, if nothing is done, **it** means that the Matignon agreements are nothing but a farce.

In (23), the author is attempting to describe his life growing up in the USA. His communicative strategy is to compare it to the kind of life described in the movie *Radio Days* by Woody Allen. This is accomplished in the first sentence by a parallel structure that establishes the entity evoked by the first *c'* (a reduced form of *ça*) as similar to that of *cela* (related to *ça*). The entity *c'* evokes is the summary of the author's life. At this point, it is fairly abstract to the hearer. The entity *cela* evokes is the summary of the events and situations presented in the film. It is readily available to the hearer, assuming she has seen the movie. The predicate that links the two entities allows the hearer to make the referent of *c'* more concrete by filling it in with elements of the movie. Later, the author makes his point more specific by establishing similar correspondences between specific elements. The entity the final *ça* evokes recapitulates and summarizes these elements.

In (24), the author is complaining about the poor treatment of the native population in New Caledonia by the French authorities following the *Matignon* agreements with France. The two examples provided are meant to be symptomatic of a general atmosphere. The entity *ça* evokes needs to be understood in that way. It cannot be said to summarize the two examples of bad treatment inflicted on the locals by the French authorities, but it generalizes them to all other possible instances of such behavior. The region profiled by *ça* is therefore quite abstract because it involves all the events that may conform to the pattern established by the preceding discourse.

In other cases, the entity evoked by *ça* is only available by an inference created by the presence of specific elements in the immediate context. This is illustrated in (25).

(25) *A Hambourg, mercredi dernier, la ville où les Beatles ont fait leur classes, McCartney a quand même réussi à conjurer par moments ce malaise étrange. On ne pourra jamais s'empêcher d'être injuste avec McCartney, de lui reprocher de ne plus être ce qu'il a été. Et c'est vrai qu'il est un peu mesquin dans le choix des morceaux : pas une chanson ne porte la patte de John Lennon,* ***ça*** *n'aurait pourtant pas manqué d'élégance, mais c'est peut-être la qualité qui lui fait le plus défaut, comme en témoigne le goût douteux des rares effets scéniques.*
In Hamburg, last Wednesday, the city where the Beatles played at their beginnings, McCartney still succeeded at times to overcome this strange discomfort. One won't ever be able to stop being unfair to McCarney, to stop blaming him for no longer being what he was. And it is true he was a little classless in his choice of pieces: not a single song with John Lennon's signature, even though **that** wouldn't have lacked elegance, but it is perhaps the quality he lacks the most, as attested by the doubtful taste of the rare scenic effects.

In (25), the pronoun's referent isn't directly available in the immediate discourse. The entity *ça* evokes is a nominal created by inference out of the preceding negative statement. It can roughly be characterized as *une chanson qui porte la patte de John Lennon* 'a song with John Lennon's signature'.

2.5. Subjective construal. One of the most important aspects of the meaning of *ça* is the subjective construal it imposes on the entity it evokes. The notion of subjectivity is used here in the sense of Langacker (1985, 1990). It refers to the vantage point from which a linguistic expression is conceptualized, as well as the viewing organization that exists between the conceptualizer and her conceptualization. The analysis of an expression's viewing organization involves the investigation of the conceptual asymmetry in the construal relation between the conceptualizing subject and the object of conceptualization. Different configurations of viewing arrangement have been shown to have important linguistic correlates (Langacker 1985, 1990, Achard 1998). The subjective construal *ça* imposes on the entity it evokes is illustrated in (26) and (27).

(26) *Les Archaos ont investi le Cirque d'Hiver, et* ***ça*** *fait du vacarme. Ils se déguisent en punks désinvoltes, pratiquent la dérision-déglingue avec une fraîcheur revigorante, en bons enfants des Monty Python et du Magic Circus, cependant respectueux des lois du cirque. Simplement, ils les habillent à leur manière.*
The Archaos have invaded the Cirque d'Hiver, and **that** makes a ruckus. They dress up as carefree punks, poke fun of everything with invigorating freshness, the rightful heirs to Monty Python and Magic Circus who nonetheless respect the rules of the circus. Simply, they dress them up in their own style.

(27) Clair de terre *n'est pas une pièce rose.* ***Ça*** *se passe dans nos campagnes.*
Clair de terre is not a light play: **That** takes place in our countryside'

The example in (26) is reminiscent of the ones in (13) and (14) in that the entity *ça* evokes represents the sub-part of reality that includes the presence of the specific proposition *Les Archaos ont investi le Cirque d'Hiver*. However, it differs from them in two major ways. Firstly, the perception of noise is presented as a most salient characteristic of reality to the point that it metonymically stands for reality. Secondly, the entity *ça* evokes is construed subjectively.

In order to understand the expression's subjective character, consider a possible alternative to (26), namely *Les Archaos ont investi le Cirque d'Hiver, et ils font du vacarme* 'the Archaos have invaded the Cirque d'Hiver, and they are making a ruckus'. In that alternative, the verb profiles the speaker's conceptualization of the relation between the source of the noise (*ils*) and their creation (*le vacarme*). The source of the process, as well as the main relation are well delineated and profiled. They are both treated as mere objects of conceptualization, while the speaker remains off stage. The subjects and objects of conceptualization are therefore maximally differentiated, but the situation is quite different in (26). The subject *ça* doesn't represent the source of the noise, but the abstract location (sub-section of reality) where it is taking place. Importantly, that abstract location also includes the speaker. The source of the noise is more diffuse and also implicitly contained in that location. The verb *faire* doesn't profile the creation of the source of the noise, but the sensation that fills the scene. The presence of *ça* therefore blurs the difference between the object of conceptualization (the source and production of the noise) and the conceptualizing subject (the speaker) by making them both part of the profiled scene.[4]

The situation is similar in (27). The application of the same kind of analysis yields the impression that the play is viewed from within, that is to say by someone actually watching the show. This characteristic of *ça* is often used to present a situation in a more lively manner, as if the conceptualizer were actually experiencing it. I will simply give one additional example in (28).

(28) *Vous avez vu un peu ce qui se passe dans ce pays ? C'est la folie !* ***Ça*** *proteste,* ***ça*** *rouscaille,* ***ça*** *rouspète dans tous les coins.*
Have you seen what's going on in the country! It's crazy! **That** protests, **that** complains, **that** bitches in every corner.

The source of social unrest in (28) is clearly people in general. A possible alternative could easily be *les gens protestent, rouscaillent, rouspètent* 'people protest, complain, bitch'. However, that source is subjectively construed, implicitly situated within the abstract setting profiled by the subject *ça*. The main verbs do not portray clearly defined actions with clearly identifiable sources, but rather the more global atmosphere that results from the subjective construal of the conceptualized scene.

3. Conclusion. This paper argued that the French demonstrative pronoun *ça* provides the hearer with the instructions to create a region within the space provided by the specific context of the utterance, and that the pronoun profiles that region. It also showed that the precise nature of the region heavily depends on the context, and that it also often involves the conceptual manipulation of information available in the discourse. Finally, it argued that the pronoun imposes a subjective construal on the entity it evokes.

---

[1] In Cognitive Grammar, the schematic value of a noun is to designate a region in some domain, a region being defined as a set of interconnected entities. Consequently, in this paper, *ça* is treated as a schematic noun.

[2] The generic character of the construction comes from the fact the category is reduced to the undifferentiated mass of the facts, beliefs, emotions it triggers in the conceptualizer. This mass construal also accounts for the fact that the construction is often considered derogatory when it describes humans. This construction, however, is very complex, and the analysis presented in this paper should only be considered exploratory.

[3] The notion of evolutionary momentum represents the dynamic aspect of reality in Langacker's Dynamic Evolutionary Model (Langacker 1991) presented in Figure 1.

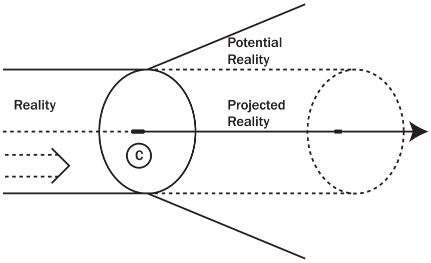


Figure 1*. Dynamic evolutionary model (from Langacker 1991:277).*

[4] One might argue that the subjective construal it imposes on the entity it evokes is a direct consequence of the pronoun's basic deictic character, but arguing in favor of this hypothesis would take us beyond the scope of this paper.

REFERENCES

Achard, Michel. 1998. *Representation of cognitive structures: Syntax and semantics of French sentential complements.* Berlin: Mouton de Gruyter.

Cadiot, Pierre. 1988. De quoi *ça* parle? A propos de la référence de *ça* pronom-sujet. *Le français moderne* 65:174–92.

Carlier, Anne. 1996. 'Les Gosses ça se lève tôt le matin': L'interprétation générique du syntagme nominal disloqué au moyen de ce ou ça. *Journal of French language studies* 6:133–62.

Grevisse, Maurice. 1986. *Le bon usage*, 12th ed. Paris: Duculot.

Langacker, Ronald W. 1985. Observations and speculations on subjectivity. In *Iconicity in syntax,* ed. by John Haiman, 109–50. Amsterdam: John Benjamins.

———. 1987a. *Foundations of cognitive grammar, vol. 1. Theoretical prerequisites.* Stanford: Stanford University Press.

———. 1987b. Nouns and verbs. *Language* 63:53–94.

———. 1990. Subjectification. *Cognitive linguistics* 1:5–38.

———. 1991. *Foundations of cognitive grammar, vol. 2. Descriptive application.* Stanford: Stanford University Press.

* * *

WHAT'S IN THE HEAD? METAPHORICAL EXPRESSIONS IN POLISH AND ENGLISH

Barbara Bacz
Laval University

Traditionally, metaphorical expressions have been regarded as *figures of speech*, i.e., as stylistic ornaments employed by writers or orators to enhance verbal performance[1]. Metaphor was a vital characteristic of what has been called 'poetic' (or 'metaphorical') language, i.e., a language containing unusual (Lakoff uses the adjective 'novel') expressions in which one or more words for a concept were used outside of their normal conventional meaning to express a concept that was somehow similar. It was widely assumed that everyday language was not metaphorical. In consequence of this view, the study of metaphor as a phenomenon was confined to the realm of literary scholarship with the rare exception of language-oriented work produced by philologists who professed an interest in stylistics.

The novelty of the Lakoff-style cognitivist approach to the phenomenon of metaphor (see the work of Lakoff and collaborators 1980–98) lies in the fact that the so-called 'metaphorical' uses of language ceased to be viewed as figurative embellishments of rhetorical style, but came to be considered (and studied) as a surface manifestation of the way in which speakers think; more precisely, of how they conceptualize abstract categories making use of their knowledge of the categories that are familiar to them from everyday experience.

One implication of the cognitivist view of metaphor for the general theory of language is, from the perspective of an ordinary linguist, that the place of this linguistic phenomenon has been shifted from the surface level of performance to the underlying level of competence. Another consequence, this time for the study of meaning, is that conceptual metaphor has come to be regarded as a phenomenon central to the semantics of everyday language, whereas traditional, poetic metaphor has been relegated to the peripheral position of conceptual metaphor's extension.

Metaphor viewed as a cognitive instrument is defined as a process of 'conceptualizing one mental domain in terms of another,' a mental operation which cognitive linguists call *a cross-domain mapping* (Lakoff 1993). The term 'domain' used in this definition should be interpreted as referring to a body of knowledge 'experienced and stored for a certain field by individuals and social groups or cultures' (Ungerer & Schmid 1996:120). Metaphorical mapping implies transfer from the source domain to the target domain. What gets transferred in the process are not just the properties of the source category but the entire structure of internal relationships characterizing the domain of the source category. Internal relationships of the source category that are exploited in a given metaphor are called *structural mappings*. They capture basic structural analogies (or correspondences) between the source and the target categories and are manifested in language by sets of specific metaphorical expressions, e.g., in the metaphor +LOVE IS A JOURNEY+, proposed and discussed by Lakoff (1992:4), one such analogy would be: THE LOVERS CORRESPOND TO TRAVELLERS, and it would be manifested by such everyday expressions as *Look how far we've come; It's been a long road,* etc. Lakoff (ibid.) insists that the principle underlying structural mapping is neither part of the grammar nor of the lexicon, but that it belongs to the conceptual system underlying the language in question (here English) which permits understanding of one domain of experience, (here love), in terms of a very different domain of experience, (here journey). One of the major objectives of Lakoff's research since 1990's has been to verify the hypothesis accounting for the principle underlying cross-domain mapping, which he calls the Invariance Principle, and which states the following: 'Metaphorical mappings preserve the cognitive topology (that is, the image-schema structure) of the source domain, in a way consistent with the inherent structure of the target domain' (Lakoff 1992: 11).

The objective of the present analysis of Polish metaphorical expressions making reference to the body-part head is to try to account for a small portion of these idioms using the cognitivist concept of metaphor as a tool of analysis. Polish metaphorical expressions related to the notion of *głowa* 'head' will be compared with English metaphorical uses exploiting the same body-part concept where the comparison appears significant in order to establish which aspects (structures) of the source category have produced similar expressions in the two languages and which ones have led to different metaphorical realizations.

Two large groups of examples relating to the concept of body-part head in the data are: (a) set lexical expressions with the word *head/głowa* applied to domains other than the domain of the body-part, and (b) idiomatic expressions in the form of sentences from everyday speech which make use of the notion of the body-part head to represent the speaker's understanding of the abstract concept of the human mind and the activity of thinking. In the following analysis, I will discuss these two types of metaphorical examples using the cognitivist concepts of *lexical* and *conceptual* metaphor, respectively.

Other metaphorical expressions, which exploit the concept of the body-part head to refer to LIFE (*Przypłacił to głową / It cost him his head*), TALENT OR PREDISPOSITION (*Ma głową do interesów / He has a head for business*), or to denote a MEASURE of length or quality (*Przerasta mnie o głową / He's head and shoulders above me*), etc. will not be considered. Also, when talking about metaphorical expressions, I am using the adjective 'metaphorical' in its broad sense meaning 'figurative', i.e., non-literal. Thus, the corpus of metaphorical expressions considered in this paper includes all kinds of figurative uses independent of their subsequent classification.

1. LEXICALIZED (OR 'DEAD') *HEAD-TO-PLANTS* METAPHORS. The cognitivist conception of metaphor as a conceptual instrument differs from the traditional perception of the phenomenon, among others, in the attitude towards linguistic units that have been

known as 'dead' metaphors. The term denotes originally metaphorical expressions, such as *head of state/ głowa państwa*, which have become so conventionalized that their metaphorical character is no longer noticed. These expressions have become lexicalized as independent dictionary entries and have thus acquired a literal meaning of their own. From a cognitive perspective, such lexicalized metaphors are not only not 'dead', but extremely important since they are the best examples of unconscious, and therefore most efficient metaphorical processes operating in language. (See Lakoff & Turner 1989:129). A typical Polish-English example of such lexicalized metaphor is given in (1):

(1) głowa/ główka [diminutive of *głowa*] *kapusty*
'a *head* of cabbage'

The conceptual metaphor which resulted in the expression *head of cabbage/ głowa kapusty* consists in the transfer of important properties and cognitive structures inherent in the source category body-part HEAD to the target category CABBAGE. The specific features and structures which underlie the metaphor, i.e., which are responsible for the transfer, are captured through the following structural correspondences:

- The shape of a head of cabbage corresponds to the shape of the human head.
- The position 'up' (above ground) of a head of cabbage corresponds to the position 'up' (with respect to the rest of the body) of the human head.

The first of the two structural correspondences distinguished for (1) is obvious: both the human head and the head of cabbage are more or less round. This feature, however, would not be sufficient to justify the presence of metaphorical mapping from the domain BODY-PART HEAD into the domains of larger superordinate categories, such as VEGETABLES or even PLANTS, a transfer which is apparently required by the Polish and English expressions containing the word *głowa/ head* such as *główka sałaty/ a head of lettuce* or *głowa słonecznika/ a sunflower head*. Beets, potatoes and radishes are also round, but their shape has obviously not evoked the association needed for a metaphorical conceptualization based on the notion of head in either Polish or English. (Polish uses the image of a bulb '*bulwa*' as the source category for potatoes, and English has *tuber*, both terms being rather technical, in spite of the fact that in both languages potatoes do have *eyes 'oczka'* [a diminutive form of *oko 'eye'*]). It seems that in the case of vegetables, size might be an important factor. Cabbage and lettuce may be about the size of a human head while potatoes, beets and radishes are normally smaller[2]. To me, however, a generalization that would apply to members of a domain larger than the domain of vegetables, i.e., the domain of plants, appears more convincing. It could be based on the observation that for the HEAD-metaphor to occur, the plant's top parts have to be seen (as the plant is growing). That's why I suggest that the second structural feature of the source category, the feature concerning the position 'up', should be considered as the most significant element in the explanation of the metaphor.

Both English and Polish exploit the concept of the body-part head metaphorically to refer to the top parts of certain flowers, such as sunflower, daisy or clover *(główka stokrotki, słonecznika / head of a daisy, a sunflower*, as in *Słonecznik zwiesił głowę* 'The sunflower has hung down its head' or *główki koniczyny / clover heads),* or generically, to the top parts of all flowers, as in *Chłopczyk poobcinał główki kwiatom w wazonie* 'The boy has cut off the heads of all the flowers in the vase'. The lexical metaphor in this case is motivated not by the size of the flowers (sunflowers are big while daisies and clover are small) but by the fact that their *heads* are round and what's more important, that they are placed 'up' on their stems (the latter feature additionally recalling the linear shape of the human body with the head on top)[3].

Discussing the use of body-part terms in expressing concepts of spatial orientation, Heine (1997:40–41) observes that 87% of all African and 61% of all Oceanic languages have grammaticalized the word for head to conceptualize the spatial orientation concept 'up'. The importance of the HEAD-to-ORIENTATION UP metaphorical mapping in the languages of the world stems from the speakers' experience of the anthropomorphic model of the vertically-oriented human body with the head on top.

Lexicalized metaphors are important for the understanding of polysemy since they represent what has been described as *multiple categorization* (See Ungerer & Schmid 1996:118). The cognitive view of metaphor permits us to interrelate categories from different domains that are expressed by the same word, and to construct complex network models for lexical items. In a possible cognitivist model of category interrelationships, the lexical item *head* will represent a subordinate category shared by two superordinate categories: BODY-PART and PLANT, and will thus demonstrate the link between these two categories established through metaphor. Also in the cognitivist framework, a semantic network model for the polysemous lexical item *head* (or *głowa*) will be constructed in such a way as to incorporate all domains which exploit the cognitive structures of the body-part prototype through the process of metaphor.

In the work of Picoche (1994, 1995), who represents the Guillaumian theory of language, lexicalized metaphors are explained through the mechanism of dematerialization. Analysing the literal and the figurative senses of the French verb *marcher*, in uses such as *L' homme marche – L'ascenseur marche*, Picoche argues that the metaphorical transfer within this lexical item results in the increased abstraction of its lexical meaning and the resultant lexical impoverishment. Explaining the metaphorical uses of the word *głowa/ head* in lexicalized expressions of the *cabbage-head* type as dematerialization might imply that the item in question is seen as being on its way to become grammaticalized (as *do* or *will* have grammaticalized in English from full lexical categories in *I will* ***do*** *it; God has* ***willed*** *it so* to grammatical categories in *I* ***do*** *intend to go; I* ***will*** *go*). The dematerialization hypothesis can perhaps be maintained if the adverbial or adjectival derivatives of the noun *head/ głowa* are brought in as evidence (such as ***headlong*** or *He is* ***ahead*** *of me* in English or *dworzec* ***główny*** *'head station' – central station*, or ***głównie*** *'mainly'* in Polish). If the word in a lexicalized metaphor does not change its part-of-speech status, however, the grammaticalization claim appears counterintuitive and premature, especially in Polish, whose grammatical system still makes productive use of diminutives or augmentatives to indicate, I would like to argue, that the word in question is used metaphorically. I suggest that the word *głowa* is used in its diminutive form in the lexical entry *główka kapusty 'cabbage head'* not to point to its small size or the speaker's affectionate attitude, but to signal to the hearer that it should not be interpreted literally[4]. On the other hand, I find the use of the base form *głowa* in the expression *głowa kapusty* marked, for it clearly implies a reference to the cabbage head's big size.

2. +*HEAD IS MIND*+ CONCEPTUAL METAPHOR IN POLISH. According to traditional rhetoric, metaphorical expressions considered in this paper represent two figures of speech: metaphor, as in (1) and metonymy, as in (2). Metaphor has been traditionally defined as a rhetorical device based on notions of 'similarity' or 'comparison' between the literal and the figurative meaning of an expression, whereas metonymy has been limited to uses based on the principle of 'contiguity' (i.e., nearness or neighbourhood) between what is denoted by the literal meaning of a word [in an expression] and its figurative counterpart (see Ungerer & Schmid 1996:115). The classical distinction between metaphor and metonymy is preserved in cognitivist theory even though in essence the two categories represent very similar processes. Both are conceptual in nature, both can be lexicalized, and both are based on the principle of source-to-target mapping. The difference between them is that whereas in metaphors the transfer occurs across different domains (from one domain to another), in metonymies, it takes place within the same domain.

(2) a. Wypada po 3 dolary *na głowę* (= na osobę).
It comes to 3 dollars *per head* (= per person).
b. Kowalski to *tęga głowa*.
'Kowalski is a big head'.
Kowalski is *very bright*; he is a brain.

The relationship between the literal and the figurative meanings of *głowa 'head'* in the Polish and English uses in (2) is typical of the classical *pars pro toto* (+ PART FOR WHOLE +) metonymy: in (2)a the word *head*, referring to a part of the human person stands for the whole person. The domain of the source and the target category is the same: the human person. In (2)b, *Kowalski to tęga głowa*, there is a combination of the + PART FOR WHOLE + metonymy (Kowalski's head stands for Kowalski) and the + HEAD IS INTELLECT + metaphor (Kowalski is equated to his intellect, and the size of his intellect is alluded to by reference to the size of his head). The Polish expression *tęga głowa 'a big head'* refers to superior intellect whereas in English *big head* suggests arrogance and conceit. Other types of metonymy distinguished in cognitivist literature are based on the contiguity relations listed in (3), after the table quoted by

Ungerer and Schmid (1996:116) on the model of the collections of metonymy types provided, among others, by Lakoff and Johnson (1980:38f):

(3) Contiguity-of-referents relations in metonymy (Ungerer & Schmid 1996:116)
+PART FOR WHOLE+ (*all hands on deck*)
+WHOLE FOR PART+ (*to fill up the car*)
+CONTAINER FOR CONTENT+ (*I'll have a glass*)
+MATERIAL FOR OBJECT+ (*a glass, an iron*)
+PRODUCER FOR PRODUCT+ (*buy a Ford*)
+PLACE FOR INSTITUTION+ (*talks between Washington and Moscow*)
+PLACE FOR EVENT+ (*Watergate changed our politics*)
+CONTROLLED FOR CONTROLLER+ (*the buses are on strike*)
+CAUSE FOR EFFECT+ (*his native tongue is German*)

From the point of view of the distinction between metaphors and metonymies, the second, sentence-based group of the Polish and English metaphorical expressions involving the notion of *głowa/head* would fall into the category of metonymy, mainly of the +CONTAINER FOR CONTENT+ type. In both English and Polish, the body-part head is conceptualized as the container for the brain, and by subsequent metaphorical-metonymic extension +BRAIN IS THE CONTAINER OF MENTAL OBJECTS+, as the seat of thoughts, ideas, wisdom and intellect; in short, as a management centre of all mental activity. Examples (4)a–e illustrate metaphorical mapping from the source domain HEAD to the target domain MIND achieved through the combination of the metonymy +THE HEAD STANDS FOR THE BRAIN+ and the metaphor +LOCUS OF THE BRAIN (HEAD) IS MENTAL ACTIVITY+. To follow the analytical procedure proposed by the cognitivists, metaphorical transfer from the domain HEAD to the domain MIND can be described by a series of specific structural mappings. Five such mappings are suggested below for Polish with representative metaphorical expressions as their justification.

(4) Structural mappings from the domain HEAD to the domain MIND in Polish:

a. HEAD IS A CONTAINER FOR WISDOM, KNOWLEDGE AND THOUGHTS
Nie *mieści mi się* to *w głowie.*
'It *overflows my head*; it's too big for my head'
It [the idea] is utterly and completely incredible; it's beyond my understanding.

On *nie ma* oleju *w głowie.*
'He *has no oil in his head*'.
He has no brains.

Ładował mu *wiedzę (łopatą) do głowy*.
'He *loaded knowledge into his head* with a shovel'.

'Kto *nie ma w głowie*, ten ma w nogach' (a proverb)
'He who *doesn't have it in his head*, he has it in his feet.'
He who doesn't think, has to run.

The container metaphor is one of the most important (Lakoff uses the term 'primary') image-schema patterns used for conceptualizing abstract concepts. The head-as-container examples (4)a constitute the largest group in the HEAD-to-MIND mapping in Polish (see Doroszewski 1980, Dunaj 1996). They have been classed together on the basis of their predicate verbs, such as *mieścić się* 'squeeze in, fit', *ładować* 'load', *mieć w* 'have in', which expressly evoke an image of a container, and because of the prevalent use of the preposition *w 'in'* with the noun *głowa*, which suggests location within. Knowledge and wisdom contained in the head are pictured metaphorically as a solid mass material that can be loaded in (with a shovel, like coal), or as a liquid (oil) that can be poured in to introduce light (cf. *Pomysł zaświtał mi w głowie* 'The idea dawned on me'). Frequently, the contents of the container HEAD is referred to in a general manner by the neuter (indefinite gender) pronoun *to 'it'*.

(4) b. HEAD IS A SEAT/ A LOCUS OF THOUGHTS, IDEAS AND WISDOM
Skąd ci to *przyszło /wpadło do głowy*?
'Where did it *come/ fall into* your *head* from?'
What put that [idea] into your head?

Wyleciało mi to *z głowy.*
'It *flew out of* my *head*'.
It has gone clean out of my head; I completely forgot about it.

Zachodzę w głowę, jak to się stało.
'I am *going deep into* my *head* to figure out how it happened'.
I am thinking very hard to figure out how it happened.

Muszę *pójś* po rozum *do głowy.*
'I have to *go to the head to fetch* some *wisdom*'.
I have to set my wits to work (to do what makes sense, to make a sensible decision)

The head is also represented as 'home' for mental objects such as thoughts and ideas. These objects are not locked in or controlled by the owner of the place at which they reside, for they are represented as things that enter (*przyjść 'come to', wpaść 'fall/ drop in')* or leave *(wylecieć z 'fly out of', wypaść z 'fall out of')* their residence. When one needs to use his wits (*rozum*), the head is the place he can fetch them (*iść po*) from[5].

(4) c. HEAD IS THE BEARER OF ONE'S PREOCCUPATIONS
Mam dzisiaj dużo spraw *na głowie*.
'I *have* many things *on* my *head* today.'
I have a lot of things to do today.

Goście *siedzą* mi *na głowie* już dwa tygodnie
'Guests have been *sitting on* my *head* for the last two weeks.'
I have been preoccupied with my guests for the last two weeks; it's a bit too much.

Nareszcie *mam* ten egzamin *z głowy*
'At last I *have* that exam *off* my head'
I have finally got rid of that exam; I don't have to worry about it

Polish idiomatic uses illustrated by the sentences in (4)c are based on the structural characteristic of the body-part head which is not exploited in the HEAD-related metaphorical expressions in English. The head in these idioms is conceptualized as that part of a person which bears all of one's preoccupations, which are viewed with a somewhat negative attitude (present and future problems, worries and all things one has to take care of). The basic linguistic proof of that interpretation derives from the semantics of the Polish locative preposition *na* 'on'. The impression produced by the expressions *mieć coś na głowie/ mieć coś z głowy* 'to have something on one's head / to have something off one's head' is that whatever is placed on the head (and usually these are concrete things denoted by concrete nouns such as *exam, guests, wedding*, etc.) is perceived as a burden one would gladly get rid of. The closest English equivalent of that mapping would be found in the idiom *to get something out of one's hair*, as in *Get it out of my hair* meaning 'You take care of it' which draws on the inherent structure of another body-part category, THE HAIR. It seems that in English the hair, and not the head, is perceived as the locus of the owner's burdensome (and therefore negative) preoccupations[6].

Other metaphorical expressions in Polish and English which exploit the domain HEAD in one language and the domain of another body-part in the other language include the English idioms which use the concept HEAD to refer to the POSITION IN THE FRONT for which Polish consistently employs the concept FOREHEAD (*at the head of the procession = na czele procesji* 'at the forehead of the procession'; *head-manager = naczelnik)*, or the various Polish idioms with the word *głowa* whose meaning is rendered in English by metaphorical mappings from the body-part domains SKIN and BACK (*nadstawiać głowę* 'to stick out one's head' = to risk one's skin; *głowa do góry* 'head up' = tails up).

(4) d. DAMAGE TO THE HEAD IS DAMAGE TO THE ACTIVITY OF THINKING
Czyś ty *upadł na głowę*?
'Have you *fallen on your head*?'
Are you crazy? Can't you think straight?

Mam dziś *urwanie głowy*.
'My *head is being torn off* today.'
I am extremely busy today.

Głowa mi *pęka*.
'My *head is cracking*.'
I have a splitting headache; my thinking is impaired.

To *marzenie ściętej głowy*.
'This is *the dream of a head that has been cut off*'.
This is utter impossibility.

The uses in (4)d provide indirect support for the specific metaphorical mapping +HEAD IS THE MEDIUM (AND PLACE) FOR/OF THE ACTIVITY OF THINKING+ by suggesting that damage to the medium of mental activity implies damage to the activity itself. When a Polish speaker wants to say that somebody's thinking is impaired or that some idea is utterly unreal, he evokes an image of a head that is cracking (*głowa peka*), that has been fallen on (*upadł na glowe*), that is being torn off (*urwanie głowy*) or even cut off (*marzenie sciętej głowy*). The examples in (4)e, on the other hand, imply that the head is perceived not only as a thinking medium but as a tool, a mechanical device that will trigger the activity of thinking when stimulated: by knocking (*postukaj się w głowę*), turning (*pokręcić głową*), moving (*rusz głową*), even breaking (*łamać sobie głowę*). The use of the instrumental case on the word *głowa* is significant for the metaphorical mapping +HEAD IS AN INSTRUMENT FOR THINKING+. As shown by the translation equivalents of the Polish examples in (4)e such as 'Don't rack your brains about it' (*Nie łam sobie tym głowy*), English exploits the domain of the body-part BRAIN to render the notion of a thinking tool via metaphor.

(4) e. HEAD IS A MEDIUM (SIMPLE INSTRUMENT) FOR THINKING
Muszę *dobrze pokręcić głową*, żeby zarobić te pieniądze.
'To make a profit I have to *turn my head* real *hard*.'
I have to think hard.

Rusz głową.
'*Set* your *head in motion*.'
Start thinking.

Postukaj się w głowę.
'*Knock yourself in the head*.'
Use your senses; be reasonable.

(4) e. Nie *łam* sobie tym *głowy*.
‘Don’t *break* your *head over* it.’
Don’t rack your brains about it.
(cf. a cross-word puzzle = *łamigłówka* ’ what breaks your head’; French *casse-tête*)

3. CONCLUSION. Lexicalized metaphors from the domain BODY-PART HEAD to the domain PLANTS are to be found in both Polish and English, and in both languages they are based on the mapping of two structural characteristics of the source category: the head’s round shape and its position ‘up’ with respect to the rest of the body.

The anthropocentric character of the source-domain BODY-PART HEAD exploited in the metaphorical rendering of our understanding of the target-domain MIND implies that the experience on which the +HEAD IS MIND+ metaphor is based is very similar across languages. According to the prediction made by Lakoff and Johnson (1980) that common experiential motivation should result in the possible universality of the metaphoric mechanisms, we should expect far-reaching similarities in HEAD-related metaphorical expressions in Polish and English following from the similarity in the specific cognitive structures that can be mapped. Sample analysis of the Polish expressions with the word *głowa* carried out here shows that the universality claim has to be taken with caution, even in the case of the relatively simple (non-cultural) metaphor which exploits the concept of the human head to talk about matters of the human mind. At least one cognitive structure +HEAD IS THE BEARER OF ONE’S PRE-OCCUPATIONS+ in the source category has been found to be present and exploited in Polish, but not in English. The reason for this divergence does not lie in the cultural differences, such as those discussed by Cienki (1999) or by Pamies and Inesta (in print) in what they call ‘cultural metaphors’, but in the differences in the conceptualization of the source concept by the speakers of the two languages. I think that the differences in the way speakers of different languages exploit cognitive structures inherent in identical source-categories (such as body-parts) are due not only to cultural factors affecting their conceptualization of these categories but are also caused by the specificity of the grammatical system they have at their disposal when using conceptual metaphor. More research in prepositional and case semantics is needed to substantiate the claim that the grammatical structure of a language (in this case, Polish) can have a bearing on metaphoric realizations in that language.

---

[1] I would like to thank two anonymous reviewers of this paper for their helpful comments and suggestions.

[2] The size factor has been pointed out to me by Bill Sullivan as a source for a possible generalization.

[3] In the Polish expression *główka czosnku ‘a head of garlic’ (a garlic bulb),* which obviously violates the position ‘up’ generalization, the shape characteristic appears so important that it overrides all other structures of the source-domain in the metaphor. In the discussion following the presentation of this paper, Stephen Straight has suggested to me that the

decomponentiality characteristic of a head of garlic, which is also present in a head of lettuce or cabbage and in the heads of sunflower, daisy or clover (all of which contain elements easily separable into smaller parts: gloves, leaves or petals) may be responsible for the presence of the same word *head* in their names. If the decomponentiality characteristic is taken as the structure responsible for metaphorical transfer from the domain BODY-PART HEAD, we would have to assume that the human head is conceptualized as a composite of its parts: the nose, the eyes, the mouth, etc. , or at least that it contains an element easily separable into more or less identical smaller units (the hair?).

[4] Additional evidence is furnished by examples such as *missile head* 'głowica [augmentative] pocisku' or *pin head* 'główka [diminutive] szpilki', etc.

[5] The English category WITS is rendered in Polish by the uncountable noun *rozum*, which can also signify BRAINS. In other idioms (e.g. *wlewać komuś olej/ rozum do głowy 'pour oil/ brains into somebody's head'*), ROZUM is conceptualized as a light-producing liquid.

[6] The English expression *to have something* **on** *one's mind,* with the preposition *on,* applies to persistent thoughts about something which can be either negative or positive, unlike the expression *to have something* **in** *one's hair*, with the preposition *in*, whose referent can only be negative. The idiom *to have something on one's mind* (as in: *You are always on my mind*) is rendered in Polish by expressions belonging to the +HEAD IS A CONTAINER OF THOUGHTS+ subcategory, with the prepositions *w* 'in' or *po* 'all over, about'. Examples: *Ciągle siedzi mi to* ***w*** *głowie* 'It [that thought] is still sitting **in** my head', with the verb *siedzieć* 'to sit' indicating persistence, and the locative preposition *w* 'in' indicating location within, or: *Ciągle mi to (ta myśl) chodzi po głowie* 'It (that thought) is constantly walking about in my head', with the imperfective verb *chodzić* 'walk' and the locative preposition *po* 'all over' referring to the area covered by the subject's movement.

REFERENCES

CIENKI, ALAN. 1999. Metaphors and cultural models as profiles and bases. In *Metaphor in cognitive linguistics*, ed. by Raymond W. Gibbs, Jr. & Gerard J. Steen, 189–203. Amsterdam: John Benjamins.

DOROSZEWSKI, WITOLD. 1980. *Słownik poprawnej polszczyzny*. Warszawa: Państwowe Wydawnictow Naukowe.

DUNAJ, BOGUSLAW. 1996. *Słownik współczesnego języka polskiego*. Warszawa: Wydawnictwo Wilga.

HEINE, BERND. 1997. *Cognitive foundations of grammar*. New York: Oxford University Press.

LAKOFF, GEORGE. 1990. The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive linguistics* 1-1:39–74.

———. 1992. The contemporary theory of metaphor, 1-45. At http://www.edu/~market/semiotic/lkof_met.html [electronic version of Lakoff 1993].

———. 1993. The contemporary theory of metaphor. In *Metaphor and thought,* ed. by Andrew Ortony, 202–51, Cambridge: Cambridge University Press.

——— & MARK JOHNSON. 1980. *Metaphors we live by.* Chicago: University of Chicago Press.

——— & Mark Turner. 1989. *More than cool reason: A field guide to poetic metaphor.* Chicago: Chicago University Press.

Pamies, Antonio & Eva Inesta. (in press). Some cross-linguistic observations on phraseology: The expression of injustice. *Acta Linguistica* 3. University Matei Bely: Barska Bystrica.

Picoche, Jaqueline. 1994. A continuous definition of polysemous items: Its basis, resources and limits. In *Continuity in semantic linguistics. Lingvisticae investigationes supplementa*, ed. by C. Fuchs & B. Victorri, 77–92. Amsterdam: John Benjamins.

———. 1995. *Etudes de lexicologie et dialectologie*. Collected by Nelly Andrieux-Reix & Geneviève Hasenohr. Paris: Conseil International de la Langue Française.

Ungerer, Friedrich & Hans-Jörg Schmid. 1996. *An introduction to cognitive linguistics.* London: Longman.

* * *

DATA AND *SCIENCE* IN INTRODUCTORY LINGUISTICS TEXTBOOKS

Douglas W. Coleman
University of Toledo

Previous studies (e.g., Botha 1970; Itkonen 1978, 1981; Yngve 1986, 1996; Coleman 1999, 2000) have demonstrated that the usual (i.e., introspective) 'data' of theoretical linguistics and the 'data' of science are quite different.

> That, and only that, which happens or obtains in space and time can be *observed… [I]ntuition*, for instance, is not a form of observation. To call all ways of gaining knowledge by the same name, i.e., 'observation,' …conceals important methodological differences. (Itkonen 1978:3)

In addition, authors in theoretical linguistics typically conflate 'examples' (abstract *types* or classes of events) with 'data' (*tokens*, or concrete events). These examples consist of cited forms which 'are typically not tokens corresponding to individual concrete observable "real-world" events' but represent an 'entire class of observation tokens' (Coleman 2000:77).

Is this state of affairs merely a historical 'accident' owing to an ongoing influence of the philosophy of language on the field of linguistics, unconsciously transmitted by experienced linguists to each generation of those new to the field? If so, then transmission will occur merely by new linguists' observation of the methodological practices of those who come before. If *not*, then we should find explicit instruction in the textbooks that serve to provide foundational training to new linguists.

I have investigated occurrences of certain key concepts in a broad range of introductory textbooks in linguistics, in the process looking at some of the same concepts as those subsumed under lexical analyses in my earlier research on theoretical and applied linguistics corpora drawn from professional journals (Coleman 1997, 1999, 2000). I focus on three major aspects in the texts: (1) what they have to say about the status of linguistics as a science, (2) what constitute(s) the object(s) of study in linguistics, and (3) what constitute the data to be used by linguistics in its examination of this object.

1. do introductory texts say linguistics is a science? Saussure very clearly identifies linguistics as a science when he says, 'Linguistics is closely related to other sciences that sometimes borrow from its data, sometimes supply it with data' (1959 [1915]:6). Shortly thereafter, he makes the point even more directly: 'In setting up the science of language within the overall study of speech, I have outlined the whole of linguistics' (ibid:17). Bloomfield (1933) refers to linguistics as 'the science of

language' (vii) and near the opening of his first chapter says, 'It is only within the last century or so that language has been studied in a scientific way…' (3). Nida makes his stand clear when he asserts that 'practical results should be based upon scientific analysis' (1976 [1949]:1, fn. 3). Whatmough refers to 'the new science of general linguistics' (1956:13), and in his glossary defines linguistics as 'the scientific study of the phenomena of language' (ibid:233); he attributes linguistics' 'closer relation to the exact sciences' to the 'application of symbolic logic and mathematics to linguistic problems' (ibid:13). Whorf (1956) describes linguistic phenomena as 'the province of the grammarian—or of the linguist, to give him his more modern name as a scientist' (211). He says other things that make it clear he views linguistics as a science, as when he refers to 'the science of linguistics' (e.g., on page 249). Whorf even compares linguistics to physics at one point (223). Crystal refers to linguistics as 'the science of language' (1987:412) and equates the terms 'linguistics' and 'linguistic science' (ibid:405). Akmajian et al. (1990:5) call linguistics 'the scientific study of human natural language'. In Gee's (1993:vii–viii) preface an extended simile of science, play, and linguistics is offered. Much later, on page 186, he makes the comparison explicit when he writes, 'linguists must, like all other scientists…' Clark et al. call linguistics 'the scientific study of language' (1994:62) and linguists 'the scientists of language' (ibid:15). Finegan calls it 'the scientific inquiry into human language' (1994:13). Macaulay (1994:3) calls linguistics 'the science of language', and O'Grady et al. refer to their topic as 'the realm of scientific linguistic analysis' (1997:xxiii). Fromkin and Rodman (1998:26) refer to 'the science of linguistics'. Radford et al. argue (1999:2) that the lack of 'controlled experiments on large numbers of subjects under laboratory conditions' does not make linguistics 'unscientific'; they also claim, for example, that the theory of UG follows 'standard scientific practice' (ibid:7). Weisler and Milekic (2000:130) ask, '[C]an you think of any other science in which subjects provide researchers with data based on intuitions?' (Indeed, we cannot, unless we include parapsychology and paraphysics, for example, among the sciences!)

Some writers are less clear on the issue. Sapir (1949 [1921]), for example, does not explicitly state his view one way or the other. He uses some of the lexicon of science when he refers to 'linguistics research' (61), the need for 'clear evidence' (103), or declares a classification scheme to be 'experimental' (140)—but such references are sporadic and few. He contrasts a linguistic analysis with a 'philosophical analysis of experience', but does so on the basis that it is 'on a sliding scale' (107), not on the basis of a philosophy-vs.-science dichotomy. Bonvillain defines linguistics merely as 'the study of interrelationships between language and other aspects of culture…' (1993:393). Jannedy et al. (1994) define a linguist noncommittally as 'someone who studies the structure of language and its use' (457). But interestingly, while the authors describe phonetics as 'an experimental science' (33), phonology, morphology, and syntax 'study' or are 'the study of' certain aspects of language (87, 129, 165); semantics 'addresses and attempts to answer' certain questions (211); and pragmatics 'concerns itself' with other things (223). Napoli refers to linguistics as a field that is 'technical' (1996:v) and 'rich, formal, exciting' (ibid:vii). She states that linguistics is defined not so much by its material as 'the way [linguists] look at data, the manner in which they argue, and the range of conclusions they find sensible in light of the data' (ibid). However, nowhere—so far as I can tell—does she explicitly identify linguistics as a science, even though this is strongly suggested by the preceding quotation. Hudson (2000) nowhere explicitly identifies linguistics as a 'science', but does contrast the approach of those working within 'generative grammar' with that of 'empiricist linguists' (499), the latter of which may be understood to work within the domain of science. Similarly, although Matthews cites a common definition found in general dictionaries—'the science of language' (1997:vii), he also immediately thereafter cites what he terms a 'more cautious wording'—'the branch of knowledge that deals with language'. As I will show, almost all of those who are in this fashion unclear or noncommittal on the scientific status of linguistics use the lexicon of science in referring to 'hypotheses', 'data', and so on.

In my entire informal sample, the authors of only one introductory text openly express doubts about the scientific status of our field. Southworth and Daswani (1974:8–9) say the drive 'to create a more empirical base for their field'

> …has not, however, been strong enough to prevent bias in data selection and presentation, or in the selection of problems for study, because of assumptions about the nature of language which are based on philosophical positions, rather than on the linguistic data alone.

2. WHAT DO THE TEXTBOOKS SAY IS THE OBJECT OF STUDY OF LINGUISTICS? Saussure (1959 [1915]) identifies the 'scope of linguistics' as '(a) to describe and trace the history of all observable languages…, (b) to determine the forces… at work in all languages…, and (c) to define and delimit itself' (6). With the exception of (c), the object of study of linguistics is clearly identified as *language*. Shortly after this, under the heading 'Definition of Language' (7), his key question is 'What is both the integral and concrete object of linguistics?' Is Saussure unaware of the problem regarding language as a 'concrete object?' No. In his discussion of the neogrammarians (5, fn. 2), he says, 'language is not an entity and exists only within speakers' and later, in this oft-quoted excerpt (8):

> Other sciences work with objects that are given in advance… but not linguistics. …[In linguistics, f]ar from it being the object that creates the viewpoint, it is the viewpoint that creates the object.

He resolves this as follows: 'One must not go too far, however, and a compromise is in order' (5) and '…*from the very outset we must put both feet on the ground of language and use language as the norm of all other manifestations of speech*' (9–emphasis his).

Others have followed his lead. Bloomfield (1933:27) calls language 'the subject of our study', clarifying the scope of linguistics by distinguishing 'the *act of speech* and the other occurrences… *practical events*' (23). In his terms, the latter are composed of 'stimuli and reactions' and are '*real*… events' (27). But to him, 'as students of language… we are concerned precisely with the speech event' (26–27). Nida defines morphology, for example, as 'the study of morphemes and their arrangements into words' (1976 [1949]). Whatmough (1956) says, 'This book is about language' (18) and identifies the objects of study in linguistics as 'the phenomena of language' (233). Whorf (1956) refers at various points to the examination or analysis of a language or of languages (e.g., on page 212). Southworth and Daswani (1974:6) explicitly identify the object of study of linguistics as '*the language itself*' (emphasis theirs). Crystal identifies language more obliquely as an 'object of fascination' but does associate its study with linguistics (1987:404). Akmajian et al. (1990:5) identify the object of study of linguistics as 'human natural language' and 'the nature of language and communication'.

Gee (1993) focuses on 'the study of language' (ix) and identifies 'a large part of what linguists study' as 'the words and syntax of particular languages and of language in general (language universals)' (13). He broadens his definition of the scope of linguistics when he says, 'linguistics also involves moving beyond words and the syntax of phrases and sentences to study how language is *used* in *real* contexts' (the emphasis on 'real' is mine) (13). He seems to see this as a positive thing, despite the implication that most of linguistics deals with *unreal* contexts. Gee here also repeats the solipsism that language is an object (like a paintbrush or a pie fork) that can be used. Those who would defend a reference to language as a 'usable object', calling it merely a metaphorical convention, should note what he says about the reality of the objects of language:

> I will call entities like cells and genes, or molecules and atoms *theoretical entities*. I call them this *not* because they don't exist (they do exist), but because they are used to *explain* the properties of things that are more readily apparent and easier to observe than they are (things like cats and water, and the sentences one hears and understands). (186)

In other words, Gee places 'sentences' on the same level of reality as cats and water!

In Gee's following paragraph (186) we see a close replay of Saussure's conceptual wrestling match with the concreteness of the objects of language, and Gee comes to basically the same conclusion as Saussure: sentences, like cats and water, are concrete.

Jannedy et al. (1994:ix) identify language as 'the object of our study'; Clark et al. (1994) describe their book as 'an introduction to the study of language' (2) and define linguistics as 'the study… of language' (62); and Finegan (1994:14) describes the object of study of linguistics as 'particular facets of languages' and 'the nature of language itself'. Macaulay (1994:3) identifies the object of study of linguistics first as 'the abstract nature of linguistic structure' and—like Gee (1993)—goes on to explain that more recently 'linguists have begun to look at ways in which people *use* language' (emphasis mine). Napoli (1996:v) identifies the subject of her book as 'the core areas of the analysis of language'. O'Grady et al. (1997:1) say that linguistics studies 'the nature and *use* of language' (emphasis mine). Matthews says linguistics is the 'science' or 'study' of language and that it 'deals with language' (1997:viii). Fromkin and

Rodman (1998:14) say linguists 'describe a language' and equate 'a true model of the speakers' linguistic capability' with 'a successful description of the grammar and of the language itself'. Radford et al. identify 'a major task for the linguist' as that of discovering 'the nature of …grammars' (which, of course, they have already said consist 'at least' of 'a lexicon, a syntactic component, a PF [phonetic form] component and an LF [logical form] component' (1999:6). Weisler and Milekic (2000:1) refer to 'the study of language' (their book is *Theory of Language*). Hudson identifies concerns of linguists as including 'how languages are structured and work, and how they are learned' as well as the application of linguistics in practical areas (2000:499).

Sapir (1949 [1921]) does not make an explicit pronouncement on the issue, but makes it clear, nevertheless, that the object of study of linguistics is language—e.g., 'the way is now cleared for a serviceable definition of language' (8) and in chapter titles such as 'The Sounds of Language', 'The Form of Language', etc.

Bonvillain (1993) is the only one of the authors from my informal sample of 22 books who disagrees. And she does not define linguistics, *per se*, as her topic, but lays out her subject matter as 'an ethnography of communication'. She identifies its components (85) in a quotation from Dell Hymes (1974:9), in which these include not only language (the 'code') but also the participants, the channel, the setting, the form or genre, and topics and attitudes of the participants.

In all of the other texts, there is no question that the object of study is language. Now, suppose we allow that 'language' as used by all these authors might simply refer to a set of properties, states, processes, what have you, in an individual. This is what Radford et al. (1999:1), for example, claim:

> [W]hen we assert that Tom is a speaker of English, we are ascribing to Tom a certain mental structure. This structure is somehow represented in Tom's brain, so we are also implicitly saying that Tom's brain is in a certain state.

This if fine, as long as we remember that we cannot observe 'mental structure'. We can posit the existence of phonemes, morphemes, sentences, etc. as what Gee (1993:186) calls 'theoretical entities' corresponding to 'mental structures', only if by that we mean we are positing theoretical entities correlated somehow with physical properties of a person. But we cannot claim to 'observe' mental structures. Sentences, for example, do not exist in the same way as cats and water.

Why is it a problem to accept language as the object of study of linguistics? Is it because elements of language are taken as theoretical entities? No. But we can see that taking elements of language as observable objects is quite unreasonable. In short, seeing language as a theory to explain observable human behavior *may* work (an empirical issue). Seeing language as a 'phenomenon' to investigate *cannot possibly* work (an ontological issue).

ABSOLUTE FORM	SANDHI-FORM
['bo:] ‘cow’	[an 'vo:] ‘the cow’
	[ar 'mo:] ‘our cow’
['uv] ‘egg’	[an 'tuv] ‘the egg’
	[na 'nuv] ‘of the eggs’
	[a ‘huv] ‘her egg’
['ba:n] ‘white’	['bo: 'va:n] ‘white cow’
['bog] ‘soft’	['ro: 'vog] ‘very soft’
['briš] ‘break’	[do 'vriš] ‘did break’

Figure 1. Examples *from Bloomfield (1933:188)*

3. WHAT CONSTITUTES DATA IN THE INTRODUCTORY TEXTBOOKS? This leads us to the question of what introductory textbooks in linguistics consider to be the data of our field.

The translator (Baskin) of Saussure (1959 [1915]) treats *data* and *evidence* as interchangeable terms, apparently referring to the same thing, for example, as ‘external evidence’ at one point and then as ‘external data’ in the next paragraph (35). In actual arguments, Saussure typically cites example forms: ‘French *au* (e.g., in *sauter* ‘jump’)’ (35); ‘Latin *crispus* ‘crisp’ provided French with the root *crép-* from which were formed the verbs *crépir* ‘rough-cast’ and *décrepir* ‘remove mortar’ ’ (83). In other words, his data, or evidence, consists of abstract, generalized *types* rather than observable *tokens* (Coleman 2000).

Bloomfield uses the term *data* repeatedly when writing in general terms, e.g., ‘All historical study of language is based upon the comparison of two or more sets of descriptive data’ (19). When presenting specific arguments, however, he uses terms like *case* (159), *form*, and *example* (160) interchangeably to refer to the material upon which he is basing his arguments. Figure 1 shows some of Bloomfield’s *examples* (188) from Modern Irish. Although at least one of these terms (*form*) would seem to exclude the gloss, it is clear that Bloomfield’s arguments work only if semantic relationships—e.g., that between ['bo:] ‘cow’ and [an 'vo:] ‘the cow’ can be assumed. Thus, Bloomfield, like Saussure before him, is working with pre-analyzed *types*, rather than observation-tokens.

Since Bloomfield seems to use *case*, *form*, and *example* interchangeably whenever he directly refers to the material that constitutes his *data*, let’s look at three instances in which he uses the term *form*. He (1933:213) says, ‘[T]he speaker who says *knives*… merely utters a ***form*** (*knives*)…’ Here, it appears that by *form*, Bloomfield refers to a communicative event (albeit hypothetical), which (if actual) would involve at least some things available for observation in a scientific sense. Later on the same page, he says a description should be arranged ‘so that as many ***forms*** as possible will be included’. This time context makes it clear he is referring to a form as a type, not an observation-token. Still, it is a type representing a hypothesized class of observation-tokens. Elsewhere on the same page, Bloomfield refers to an *underlying form* in much the same sense in which it has been used more recently: 'an indefinite number of plural nouns retains the [-s] of the underlying ***form***'. This use of *form* clearly refers to a type which exists only as a theoretical entity, and which therefore cannot be directly observed, so it has no direct relationship to observation-tokens. Such confusions of types and tokens are pervasive throughout Bloomfield's *Language*, just as they are through later textbooks (just as they are through journal articles in theoretical linguistics [TL]—Coleman 2000).

Sapir's (1949 [1921]) usage is very similar to Bloomfield's. In the body of his arguments, he generally refers to his material as *forms*, *examples*, or even as *types*. When speaking about method, however, as in footnote 7 on p. 212, he at least sometimes uses the term *data*: 'By working back from such *data* as we may possess…' In all cases, his forms/examples/types/data are, however, types, not actual observation-tokens.

Much of Nida's (1976 [1949]) *data* may have been obtained from interaction with informants. But like the *data* of Coleman's (2000) TL corpus, some is from secondary sources, and it is not made clear which is from naive informants, which from descriptive grammars, which from scholarly articles — and therefore which may actually be invented material. This is not an empty charge based on an assumption of guilt until proof of innocence. Consider Nida's use of the term *data* in the following instructions to an exercise on some homophonous forms in English (58):

> If the forms are related in meaning, determine whether the forms are in complementary distribution. (Supplementary *data* must be supplied by the student.)

Here, Nida invites the student to invent his/her own data.

Whatmough (1956) refers at least once to a *sample* (184) in a way that is somewhat consistent with usage in the social sciences and with that of qualitative analyses in current corpus linguistics. Generally, however, he cites types, but without referring to them as *data*, *forms*, *cases*, etc., e.g., 'and likewise in French *deux cents* and *cens deux*' (136).

Whorf (1956; e.g., 45, 52, 124, 168) most often refers to his material as *examples*, *samples*, or *forms*—in almost all cases, they are types, not observation-tokens. However, there are instances in which he makes reference to material actually observed, e.g., in his discussion of a Mayan inscription in the Temple of the Tepoztecatl, of which he provides an annotated sketch (44). In discussing the inscription, he alternately refers to its elements with the terms *form*, *sign*, *hieroglyph*, *hieroglyphic element*, *grapheme*, *figure*, and *glyph*. While some of these (e.g., *form*) may arguably refer to observable visual data in a scientific sense, others refer unambiguously not to observable objects, but to inferred types (e.g., *grapheme*). Yet they seem to interchangeably refer to the same elements in the illustration. Is *this* just a careless use of terminology, or is it indicative of a genuine confusion? Consider what Whorf (223) says elsewhere:

> Physics does not begin with atomic structures and cosmic rays, but with motions of ordinary gross physical objects and symbolic… expressions for these movements. Linguistics *likewise* [see below] does not begin with meaning nor with the

> structure of logical propositions, but with the obligatory patterns made by the gross audible sounds of a given language and with certain symbolic expressions of its own for these patterns.

But then try to put this together with the following (248):

> First, the plane 'below' the strictly linguistic phenomena is a physical, acoustic one, phenomena wrought of sound waves.

If 'strictly linguistic phenomena' are not physical phenomena, then the *likewise* in the earlier quotation is cancelled out. We see a genuine confusion of physical and non-physical (objective and subjective) domains here.

This confusion of tokens (*data*) and types (*examples*) is pervasive throughout almost all the textbooks, mirroring that observed in the TL corpus of Coleman (1999, 2000). Southworth and Daswani (1974:149–50) call some of the same things *examples* at one point and *data* at another (e.g., 150, 170). Akmajian et al. (1990) also do this (e.g., 119) and at other points use the terms interchangeably (e.g., 186, 188). In developing one analysis, Gee (1993) calls his items (21)a–(23)b *examples* and his (25)a–(27)b *data*. Jannedy et al. (1994) typically call the material of their exercises *data*, but occasionally say it consists of *forms* (e.g., 155) or *examples* (e.g., 207). Some authors in Clark et al. (1994) seem to prefer the term *data* over *example*, or vice versa, and the confusion arises overtly here and there (e.g., 323–24, where the same items are called *examples*, then *data*). Finegan (1994) calls material for analysis in parallel exercises *examples*, *forms*, or *data* (e.g., 72–73). Macaulay (1994) generally avoids the term *data*, preferring to call his material *examples* (and occasionally *forms*); the same confusion of types and tokens is evident, however, when we compare references to cited forms (types) as *examples* (e.g., 154, 158) and elsewhere, actual material from recorded interviews (tokens) as *examples* (e.g., 119–20). Napoli (1996) refers to the same material as *data* and *examples* (e.g., 129). O'Grady et al. (1997) call two sets of parallel items in the same analysis *data* and *examples*, respectively (69). In Fromkin and Rodman (1998), the same kind of material is referred to as *examples* in one place and as *data* in another (cf. the Russian *examples* on 72 and the Paku *data* on 96). In Radford et al. (1999:2), a discussion of *data* in linguistics includes the use of the following, referred to as *examples*: *The dog chased the cat* and **Cat the dog chased the*. In Weisler and Milekic (2000:130–31), the same material is called *some examples* and then *the previous data set*. Hudson (2000) calls material for analysis in exercises *data*, but alternates between the terms *data* and *example* when presenting analyses himself (cf. an analysis of Amharic verbs 65 ff. and one of Spanish verbs 77 ff.).

I could not find a single reference to *data* in Matthews (1997); he seems to refer exclusively to *forms* or *examples*, and in fact all his *forms* and *examples* are types, not tokens. He does avoid the confusion, by avoiding mention of any data.

Most of the texts briefly mention the availability of forms of *data* other than that invented by the researcher, but provide only lip service in this regard. Crystal

(1987:esp. 410 ff.) is one who presents a broader view of data sources in linguistics, including not only information provided by informants, but experiments, corpora, etc. Another is Bonvillain (1993), who very consistently refers to linguistic types as *examples* and observation-tokens (properties of people, the setting, etc.) as *data*, except in her introduction (4), where she refers to all types of evidence as *data*. Except for the one case in the introduction, her usage parallels that of applied linguistics, as found in Coleman (1999, 2000).

4. CONCLUDING REMARKS. Linguistics calls itself a science. Linguists clearly aspire to be scientific.

The object of study in any natural science is—to use Saussure's terms—'given in advance'. It is not created by the viewpoint of the observer. And this is even true of the social sciences, such as applied linguistics, in which observations are typically of people—their properties, and their actions, as shown by the corpus-based research of Coleman (1999, 2000). Theoretical linguistics can accurately call itself a 'science' only if it starts framing its theories in terms of real-world ('given-in-advance') objects such as those of applied linguistics or those mentioned by ethnolinguists such as Hymes (1974) or—within the field of general linguistics—Yngve (1986, 1996).

I have demonstrated the pervasive confusions in introductory linguistic textbooks with regard to what constitutes linguistic data—confusions between observation-tokens and inferred types. This is a specific instance of what Yngve (1996:24, 33) calls *domain confusions*, confusions between the physical and logical domains. We are teaching our students to be confused. Why? The linguistic community as a whole seems to be confused on this issue.

We need to rely on predictive power as our primary criterion for acceptance of a theory in linguistics. Now, is this merely a call for us to seek what has been for years called 'descriptive adequacy?' No it is not. The predictive power of a model can be measured only in terms of its 'fit' with data. The pervasive confusion in linguistics over whether scientific data can consist of observation-tokens (which it can) or inferred generalized types (which it cannot) prevents us from being able to measure the predictive power of competing models.

In short, what we need to do is to begin by scrapping all notions of 'data' as something the researcher can invent, for that matter, as anything existing in the non-physical domain. The fact that we lack physical theories of cause-effect on which to base our claims *may* excuse us for positing theoretical entities like phonemes, morphemes, and so on, entities which *may* have physical correlates in the brain. But it does not justify our using as 'data' things that we know very well we are not 'observing' in any scientific sense.

REFERENCES

Akmajian, Adrian, Richard A. Demers, Ann K. Farmer & Robert M. Harnish. 1990. *An introduction to language and communication*, 3rd ed. Cambridge MA: MIT Press.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart, and Winston.

Bonvillain, Nancy. 1993. *Language, culture, and communication: The meaning of messages*. Englewood Cliffs NJ: Prentice-Hall Publishers.

Botha, Rudolph P. 1970. *The methodological status of grammatical argumentation*. The Hague: Mouton.

Clark, Virginia P., Paul A. Eschholz & Alfred F. Rosa (eds.). 1994. *Language: Introductory readings*, 5th ed. New York: St. Martin's Press.

Coleman, Douglas W. 1997. Science or philosophy? Progress report on a lexical analysis of corpora from two journals in linguistics. In *PALC '97: Practical applications in language corpora*, ed. by B. Lewandowska-Tomaszczyk & J. P. Melia. Łódź, Poland: University of Łódź.

———. 1999. Assumptions, hypotheses, and theories in 'applied' and 'theoretical' linguistics. *LACUS forum* 25:461–72.

———. 2000. Linguistic data and its uses: The problem of conflating 'data' and 'examples'. *LACUS forum* 26:71–80.

Coulmas, Florian (ed.). 1981. *A Festschrift for native speaker*. The Hague: Mouton.

Crystal, David. 1987. *The Cambridge encyclopedia of language*. Cambridge: Cambridge University Press.

Finegan, Edward. 1994. *Language: Its structure and use*, 2nd ed. Fort Worth TX: Harcourt Brace College Publishers.

Fromkin, Victoria & Robert Rodman. 1998. *An introduction to language*, 6th ed. Fort Worth TX: Harcourt Brace College Publishers.

Gee, James Paul. 1993. *An introduction to human language: Fundamental concepts in linguistics*. Englewood Cliffs NJ: Prentice-Hall Publishers.

Hudson, Grover. 2000. *Essential introductory linguistics*. Malden MA: Blackwell Publishers.

Hymes, Dell. 1974. *Foundations in sociolinguistics: An ethnographic approach.* Philadelphia: University of Pennsylvania Press.

Itkonen, Esa. 1978. *Grammatical theory and metascience: A critical investigation into the methodological and philosophical foundations of 'autonomous' linguistics.* Amsterdam: John Benjamins.

———. 1981. The concept of linguistic intuition. In Coulmas 1981:127–40.

Jannedy, Stefanie, Robert Poletto & Tracey L. Weldon (eds.). 1994. *Language files: Materials for an introduction to language and linguistics*, 6th ed. Columbus OH: Ohio State University Press.

Macaulay, Ronald. 1994. *The social art: Language and its uses.* New York: Oxford University Press.

Matthews, Peter. 1997. *The concise Oxford dictionary of linguistics*. New York: Oxford University Press.

Napoli, Donna Jo. 1996. *Linguistics*. New York: Oxford University Press.

Nida, Eugene. 1976. *Morphology: The descriptive analysis of words*, 2nd ed. [Copyright renewed — originally published in 1949.] Ann Arbor MI: University of Michigan Press.

O'Grady, William, Michael Dobrovolsky & Mark Araonoff. 1997. *Contemporary linguistics: An introduction*, 3rd ed. New York: St. Martin's Press.

Radford, Andrew, Martin Atkinson, David Britain, Harald Clahsen & Andrew Spencer. 1999. *Linguistics: An introduction*. Cambridge: Cambridge University Press.

Sapir, Edward. 1949. *Language: An introduction to the study of speech*. [Copyright renewed — originally published in 1921.] New York: Harcourt, Brace & World.

de Saussure, Ferdinand. 1959. *Course in general linguistics*, translated by Wade Baskin. [First edition, in French, was published in 1915.] New York: Philosophical Library.

Southworth, Franklin & C. J. Daswani. 1974. *Foundations of linguistics*. New York: The Free Press.

Weisler, Steven E. & Slavko Milekic. 2000. *Theory of language*. Cambridge MA: MIT Press.

Whatmough, Joshua. 1956. *Language: A modern synthesis.* New York: The New American Library.

Whorf, Benjamin Lee. 1956. *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*, ed. by John B. Carroll. Cambridge MA: MIT Press.

Yngve, Victor H. 1986. *Linguistics as a science*. Bloomington IN: Indiana University Press.

———. 1996. *From grammar to science: New foundations for general linguistics.* Amsterdam: John Benjamins.

* * *

VERBS OF TOLERANCE WITH THE GERUND AND THE INFINITIVE

Patrick J. Duffley
Université Laval

While tolerance has achieved recognition as one of the cardinal virtues of political correctness, it has not yet been the subject of much discussion among English grammarians. The verbs which express this notion in English—*tolerate*, *bear*, *stand*, *abide* and *endure*[1]—do however present a number of points which are of interest for the grammarian. Most of these verbs can be construed either with the infinitive or the gerund. Curiously however, as noted by Visser (1973:1877), 'a semantic difference between the two alternatives—one with the infinitive, the other with the form in *-ing*—is not discernible'. This is not always the case with verbs which admit these two complement forms, as there is often a clear semantic distinction between the two constructions; cf. the examples in (1)–(3):

(1) a. I tried to open the window. b. I tried opening the window.
(2) a. I remembered to call her. b. I remembered calling her.
(3) a. I regret to tell you about this. b. I regret telling you about this.

The apparent absence of a clear semantic difference between the gerundive and infinitival constructions with verbs expressing the notion of tolerance is thus an aspect of usage which needs to be explored. Are there any cases where a distinction is perceivable? Can any explanation be offered for the closeness in meaning between the two constructions?

Another point of interest is the absence of any attested use of the *to*-infinitive after the verb *tolerate* itself. Infinitival constructions are attested with the verbs *bear* (4), *stand* (5), *abide* (6) and *endure* (7):

(4) …he thrust one hand into his trousers pocket and pulled out the six slugs he had taken from the revolver. (…) He put the shells on the table, as though he could no longer bear to hold them. (BUC L18 0770)
(5) I cannot stand to stay here any longer. (Visser 1969:1326)
(6) You know I cannot abide to light my Woodbine from the campfire. (BNC HWN 1538)
(7) But Carteret's wife, who frequented health spas, could not endure to live with him or he with her: there were no children. (BNC CD2 1061)

In no case to my knowledge however is the verb *tolerate* construed with the infinitive, a fact for which an explanation will be offered in this paper.

A third curiosity concerning the verbs of tolerance also has to do with the infinitival construction. All of the cases of this construction which have been attested occur in nonassertive contexts (cf. Quirk et al. 1985: 83–85, 784–85) where the verb of toleration is preceded by the modal auxiliary *can* with the latter being either negated or questioned, as illustrated in (8) and (9):

(8) … Stuart had such a happy, earnest look of proud possession on his face that Maggie couldn't bear to do anything to quench it. (BUC P01 1690)
(9) Beats me how she can bear to enter their hovels! (LOB P11 43)

Included under the general notion of nonassertiveness are contexts containing semi-negatives (10), restrictives (11) and negative polarity items (12):

(10) I can scarcely bear to leave you behind even for a few weeks. (LOB P13 88)
(11) He could only bear to have a doctor who did not take the absurd business too seriously, who realized that the whole sad joke of men living for only seventy years was made worse if you treated it as anything else. (LOB K06 54)
(12) He found himself more and more surprised that any person could bear to be hard to another. (LOB K06 132)

In contrast to the infinitive, the gerundive construction does occur in assertive constructions, as in:

(13) As regards simple absence of humour, I like to think I'm as fond of a good laugh as the next man, but I can stand doing without for long periods when reading, having been trained in the Oxford English school… (LOB G36 131)
(14) Here Mrs. Lowyer endured being alternately frozen and scalded by a terrified junior, then her hair wound around rollers and her scalp speared with plastic pins, and the torture was finally rounded off by being put under a red-hot dryer which she was unable to make cooler, however much she turned the knob on the dial. (BNC FA5 2231)
(15) Such fibrous-rooted subjects as bedding plants, lettuces and most herbaceous perennials will tolerate being transplanted. (BNC ACX 1586)

Other aspects of usage could be mentioned, but those referred to above will present a sufficient challenge for our explanatory hypothesis.

This hypothesis holds that the explanation of complement usage of the gerund and the infinitive requires taking into account the following three parameters:

a. the semantics of the verbal complement itself, i.e. the *-ing* form or the *to*-infinitive

b. the function of the verbal complement with respect to the main verb
c. the lexical meaning of the main verb.

Regarding the semantics of the verbal complement, the *-ing* form will be assumed to have its 'gerundive' value in usage with the verbs of toleration[2]. In this actualization of the *-ing*, the aspectual value of imperfectivity associated with the *-ing* in the progressive construction is not present. This can be seen from the compatibility of the gerundive *-ing* with the preposition *after* as compared to the impossibility of the progressive:

(16) a. *After he was eating his hamburger, he went home.
b. After eating his hamburger, he went home.

The gerundive *-ing* evokes an event as a whole much like a noun does (cf. Duffley 1995:8-10); compare (16)b above with (16)c below:

(16) c. After the meal, he went home.

The *-ing*'s function with the verbs of toleration is also like that of a noun: it simply performs the role of direct object with these verbs. This can be seen from the possibility of replacing it by a nominal pronoun:

(17) a. I can't stand seeing you lick the boots of a man like Willi. (Visser 1973:1877)
b. I can't stand it.

and the possibility of passive constructions, which show that the *-ing* corresponds semantically to that which is 'X ed':

(18) a. Some people can stand living alone for long periods of time.
b. Living alone can be stood by some people for long periods of time.

In contrast, the *to*-infinitive does not allow passive constructions (19), nor substitution by nominal pronouns, the anaphoric form corresponding to the infinitive being a stranded *to* (20):

(19) a. Some people can stand to live alone for long periods of time.
b. ?*To live alone can be stood by some people for long periods of time.
(20) a. I cannot stand to stay here any longer.
b. … and they can't stand to either.
c. *… and they can't stand it either.

It will be held here that the infinitive is a prepositional phrase whose relation to the main verb is defined by the preposition *to*. The latter represents the infinitive's event as the endpoint of a movement leading from the point in time corresponding to the main verb's event up to that of the infinitive (cf. Duffley 1992:16–18). According to whether this movement is understood to be realized or not, the infinitive will be evoked as the goal or result of the main verb's event. In either case, it is represented as subsequent to the latter.

This leaves the third and crucial factor—the lexical meaning of the main verb. Two of the five verbs under study involve the notion of holding up under some sort of strain or pressure without yielding: *bear* and *stand*. *Bear* is based on the metaphor of supporting a weight and is defined by Webster's (1966) as 'sustaining an affliction, onerous or difficult, without breaking or flinching'; *stand* is based on the metaphor of maintaining an upright position in the face of attack or aggression and is defined as 'to resist without yielding or retreating'. Two other verbs involve the idea of making it through the duration of a trial without succumbing: *endure* and *abide*. The latter also has overtones of overcoming impatience or disgust (cf. *I cannot abide such people*; *I can't abide the taste of caraway*). Lastly, *tolerate* is defined as 'to endure with forbearance and restraint' and evokes the idea of not doing anything to combat or eliminate a source of affliction or disturbance.

A further grouping may be made of the first four verbs, which contain an element of meaning which is not present in the last one. *Bear*, *stand*, *endure* and *abide* all involve an idea of coming through a trial or difficulty successfully: *bear* and *stand* denote a holding up under strain or attack without giving way or flinching; *endure* and *abide* evoke making it through the duration of a trying situation without succumbing to weariness, impatience or disgust. This element of meaning, it will be claimed here, is what explains the construal of these verbs with *can* and the *to*-infinitive. The notion of facing a trial successfully has a natural connection to the idea of whether someone has what it takes to hold up without succumbing, i.e. the capacity to bear, stand, endure or abide. From the point of view of the capacity of facing a trial, it is also quite natural to conceive the trial itself as something subsequent to this capacity. Since the trial or difficulty to be faced is expressed by the infinitive, it is understandable that the latter should be represented by means of the preposition *to* as an after-position with respect to the capacity of facing it. Thus a construction such as (21) means 'he did not have the capacity required to say an unkind word to anyone':

(21) He could not bear to say an unkind word to anyone. (LOB C08 76)

The fact that all of the uses of the *to*-infinitive occur in negative or nonassertive contexts also lends support to our hypothesis: the notion of a lack of what is required to get through the infinitive's event successfully or a doubt as to whether one has what it takes to hold up under it makes the before/after relation between the main verb and the infinitive particularly salient.

In contrast, the three cases of the gerund construction observed thus far with *can* + negation all evoke existent situations:

(17) a. I can't stand seeing you lick the boots of a man like Willi. (Visser 1973:1877)

(22) If it hadn't been her dad's week for night shift, her mum wouldn't have minded. As it was, Mum couldn't stand being alone in the house at night… (LOB G36 131)

(23) … John had drifted away from the gang. Mousie said it was because he was too proud to stand pity. Others thought he couldn't stand seeing Linda, Mrs. Bobbie Evans, still so beautiful, so much in command of everything. (BUC P17 0610)

Here it is a lack of the capacity to endure what actually exists or has existed which is evoked. It is noteworthy that the *-ing* could not be used in a context such as (21) above where the complement's event is understood to be non-realized.

This is not to say that the *to*-infinitive cannot be used to refer to existing events. Such is the case in:

(24) Do please tell Pepita. I can't bear to see her like this. (LOB P08 106)

However the infinitive is associated with the future-oriented impression of 'any longer; any more' in this type of use. This notion is even sometimes explicitly expressed as in (4) and (5) above. The infinitive is not possible however in contexts where the reference is to the actual sustaining of an actualized event. This is the case in (14), where the verb *endure* is not preceded by a modal and occurs in the past tense. Here there is no before/after relation between the enduring and the being alternately frozen and scalded; the *-ing* complement evokes simply that which was endured by Mrs. Lowyer.

Turning to the verb *tolerate*, we can now offer an explanation for the fact that this verb is not construed with the *to*-infinitive. Since it does not involve the notion of success in coming through a trial or difficulty (but rather the idea of putting up with a situation) it does not pose the question of whether the subject has what is required to hold up under a trial without succumbing. Consequently the difficult situation is represented as that which is or cannot be tolerated and not as something which the subject lacks the capacity to sustain without giving way. This can be illustrated by (25):

(25) Rose no longer wished to go out, nor could she tolerate being left on her own, so she became totally dependent on other people. (BNC ASO 1168)

However since in many cases the inability to tolerate something logically implies the lack of the capacity to sustain it without losing one's patience, a sentence such as (25) is very close, as far as the overall message conveyed is concerned, to the construction with the *to*-infinitive *nor could she bear to be left alone*. It is the choice of the linguistic means for conveying the message—in this case the choice of the lexical notion of 'tolerating' rather than that of 'bearing'—which situates the speaker mentally on the level of not putting up with the event itself, rather than on the level of lacking the capacity to perform it without giving way or losing one's patience, a view which allows a construal of the absence of the capacity as a before-position with respect to the being left alone.

As can be seen from the remarks above, we are in a very delicate area of English grammar where the lexical notions are closely related to one another and difficult to tease apart. We hope to have shown nevertheless that English usage is not haphazard with verbs of toleration and that the special behaviour of the verb *tolerate* is due to its lexical content being slightly different from the other members of this tiny semantic field. It remains to be seen whether our hypothesis will stand the test of further scrutiny and whether the insights gleaned from this examination of verbs of toleration can be related to complementation constructions with the infinitive and the *-ing* form after other types of verbs.

---

[1] The construction put up with is not considered in this study as it is not a simple verb but a complex phrase made up of verb + particle + preposition. Its exclusive construal with the *-ing* form seems attributable to the incompatibility between the preposition with and the *to*-infinitive phrase, an incompatibility which is characteristic of all prepositions.

[2] According to the traditional view, the *-ing* is held to be a 'gerund' when it has the function of a noun in the sentence (cf. Quirk et al. 1985: 1292). Such is the case in a use such as *Painting a child is difficult*, where the *-ing* form fulfils the role of subject. For a discussion of the semantic relation between the 'gerundive' and 'participial' uses of the *-ing* see Duffley 1995.

REFERENCES

BNC = British National Corpus (on-line).
BUC = W. Nelson Francis & Henry Kučera (eds). 1964. *Brown University corpus of American English*. Providence RI: Department of Linguistics, Brown University.
LOB = Stig Johansson, Geoffrey N. Leech & Helen Goodluck. 1978. *Lancaster-Oslo/Bergen corpus of British English*. Oslo: Department of English, University of Oslo.

~ ~ ~

Duffley, Patrick J. 1992. *The English infinitive*. London: Longman.
———. 1995. Defining the potential meaning of the *-ing* form in English. *Langues et linguistique* 21:1–11.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
Visser, F. Th. 1969. *An historical syntax of the English language*, vol. 3, first edition. Leiden: E.J. Brill.
———. 1973. *An historical syntax of the English language*, vol. 3, second edition. Leiden: E.J. Brill.
Webster's third new international dictionary. 1966. Springfield MA: G & C. Merriam Co.

~

FASHIONABLE *YOU*

Connie Eble
University of North Carolina at Chapel Hill

In the last few years I have been exploring the notion of fashion in lexical choice. In the process of trying to be alert to vogue expressions, I have been hearing and seeing in print the pronoun *you* not directed to a specific addressee, as in (1) and (2):

(1) You have the KLA attacking some of the French troops. (Georgie Anne Geyer, *Washington Week in Review*, 20 August 1999)[1]

(2) You did not stay in the Commodore for the quality of its rooms. The only thing that came with your room at the Commodore was a 16 per cent service charge, and whatever you found in the blue-and-gold shag rugs. (Thomas Friedman, *From Beirut to Jerusalem,* 1989)

After thirty years of teaching English composition, my immediate reaction is to grab my green pen and caution, 'Use precise vocabulary. *You* is a second person pronoun directed to an addressee. You are not addressing a speaker/reader here. Revise this sentence'.

What really needs to be revised is my own understanding of *you*. The sheer volume of examples I keep encountering in my sporadic efforts at observation attests that in ordinary, clear, spoken English and even in careful writing, the pronoun *you* is simply not being confined to addressing a specific audience of one or more. This paper gives examples of extended functions of *you* in current American English and interprets them as an effect of the movement towards informality in American society during the twentieth century.

The non-addressing function of the second person pronoun (for convenience sometimes called *indefinite you*) has been part of English usage for more than four centuries. Most reference books about the English language do include indefinite *you*, but usually not in detail and often merely to characterize it as inappropriate in writing because of its colloquial or informal flavor.

Two of the handbooks most widely used in American colleges and universities—*The Allyn and Bacon handbook*, 3rd ed. (1992) and *The Harbrace college handbook*, 13th ed. (1998)—acknowledge that *you* can have an indefinite referent, but both caution against this use in 'academic' or 'very formal writing'.

In a cursory check of several popular guides to correct usage in writing, I found that most are silent on the subject. Their major concern with the second person pronoun is to distinguish the spellings *your* and *you're*. Some of these mention *you* with indefinite reference in an entry on indefinite *one*, in which *one* is recommended as the preferred and more formal substitute for *you*. Bryan Garner's brand new *The Oxford dictionary of American usage and style* (2000) does not have an entry for either indefinite *you* or indefinite *one* but has an entry distinguishing *your* and *you're* that uses indefinite *you* (356): '"In boxing you don't have that kind of luxury or time. If you mess up in a fight or two *you're* [read *your*] career could be over". (*Lancaster New Era* [quoting an interviewee])'

By far the best treatment in a usage manual—as you might guess—is the entry for *you* in *Merriam-Webster's dictionary of English usage* (1989). It explains the indefinite *you* as related to the addressing function of the second person and points out the 'informal effect' using five citations from oral and written sources from 1957 through 1987, including one from an interview with Nobel Prize winner William Faulkner. The Merriam-Webster reference volume also links the development of *you* to 'the effect social forces have on the language' (970).

Dictionary treatments of indefinite *you* are brief. The *Oxford English dictionary,* 2nd ed. (1989: xx, 766) documents its use in seven citations dated from 1577 through 1870. *Webster's third* (1966 [1961]: 2654) defines it with the synonym *one* and gives one citation. More pertinent, however, is the subentry under *you* for the phrase *for you*, defined as 'displayed in his, her, its, or their characteristic nature or behavior', with the citation 'he's already forgotten he's hurt—that's a child for you'. This is clearly a *you* lacking a specific addressee, and it is clearly not an informal variant of *one*. *Merriam-Webster's collegiate dictionary*, 10th ed. (1993: 1374) simply gives the synonym *one* with no citation or commentary. *American Heritage* 3rd ed. (1992: 2072) says, 'Used to refer to an indefinitely specified person—one: "You can't win them all" '.

Standard reference grammars all mention *you* with indefinite reference but invariably give more inches to indefinite *one*. The best treatment of *you* in a grammar book that I have found is in Douglas Biber et al's *Longman grammar of spoken and written English* (1999), which is based on a corpus of more than forty million words, both spoken and written, divided into four core registers: conversation, fiction, news, and academic prose. In distinguishing *one* from indefinite uses of *we, you,* and *they*, the grammar observes that the *personal* pronouns 'tend to retain a tinge of their basic meaning' (331). Furthermore, the treatment of comment clauses in Biber et al gives examples of finite clauses with indefinite *you* from the news and academic prose registers that show 'recognition of the audience's knowledge or expectations' [See, for example, *as you might guess* three paragraphs above in this piece of academic prose.]. 'In news and academic prose, the use of the subject *you* in these clauses creates a more involved text, overtly attributing an attitude to the reader' (866). I find this an accurate insight into the effect of *you* by Biber et al.

However, aside from comments here and there, the semantics of *you* has not been the focus of much analysis. In 1978, M. Stanley Whitley pointed out the shortcomings of a syntactic feature analysis of *we, you,* and *they* and proposed to overcome them by a series of constraints. In response, Dwight Bolinger (1979) offered instead an interpretation based on the process of metaphor. Bolinger pictures the extension of *you* as a kind of metaphoric transfer along two continua—specific to generic and personal to impersonal. *You* allows speakers both to generalize and to personalize. Says Bolinger, 'The interpretation is open, and goes the way that best fits the pragmatics of the sentence' (208). Furthermore, he finds *you* to be a courtesy device, a way of inviting another to share a viewpoint (205)—which seems to be what is going on in the comment clauses cited in Biber et al.

The reigning expert on personal pronouns in English is Katie Wales, author of *Personal pronouns in present-day English* (1996). The cutoff date of her book is 1995. She collected data from more than twenty grammar books of the twentieth century and also searched the corpora of English usage at University College London and others.

Wales anchors her treatment of *you* in the concept of *deixis*. The first and second person pronouns are fundamentally deictic (as opposed to third person, which is anaphoric), for their reference is determined in relation to a specific speaker and addressee in a specific time and place of utterance. First and second person pronoun referents are thus not fixed or stable but shift according to the situation, as participants take turns to speak. The choice of first person pronoun or second person pronoun is essentially context-dependent. The context can be either the immediate situation (specific exophoric reference) or the larger cultural context or presupposed knowledge (homophoric). Such instability of reference is what has allowed *you* to expand into the numerous and varied acts of language in which it now occurs.

As the examples that follow indicate, indefinite *you* in contemporary American English occurs in both fiction and nonfiction writing to convey the feeling of immediacy associated with spoken English. In spoken English, indefinite *you* allows for varying degrees of generalization and personalization, and sometimes the referent of *you* is actually the speaker (first person) not the addressee (second person). *You* is often used in situations in which a speaker conveys new information by appealing to a common or reasonable bond of experience between the speaker or topic and the hearer. Establishing such rapport seems to be a device in particular of academic discourse in order to encourage engagement in learning, and *you* is frequently observable in explanations of abstract and hypothetical subject matter. *You* that applies to 'everyone' is often used in aphorisms. The indefiniteness of *you* is complete when *you* occurs as the grammatical agent of existential verbs like *have* and *got,* yet the use of *you* somehow involves the hearer in the situation.

Probably the main reason that *you* with indefinite reference has received little analysis is that it seldom occurred in writing. The favored point-of-view of writing with a serious purpose has traditionally been the third person, which distances the writer and reader from the subject matter. A third person perspective gives an air of objectivity and removal from the participants in the communication. It is the person most suitable for transmitting facts—one of the major aims of written communication since the Renaissance. In the formal written register in particular, the tone of objectivity was considered compromised by use of first and second person pronouns. The formal written register was the school model for correct writing, to the extent that writers of academic essays customarily referred to themselves in third person with phrases like *the present writer*. From their earliest days of writing instruction, writers developed a style that avoided the first and second person pronouns except in personal writing. The infelicitous *I*, *we*, or *you* that did escape an author's pen was dutifully excised by a careful editor. It is only recently that *you* has been a feature of written English. Its increased frequency corresponds to positive associations with informality that developed during the twentieth century.

Part of the modern feel of twentieth century fiction comes from a point-of-view that connects the narrator or main character with the reader and gives the feeling that the character in the text is in conversation with the reader. The ambiguous use of *you* (as direct address to a particular reader or readers and as a generalized 'any person' that will include any potential reader) is sometimes one vehicle for that connection.[2]

In *To have and have not* (1937), Ernest Hemingway uses *you* as a framing device, beginning with an ambiguous, questioning *you* that is an address to a reader who is 'anyone' and who is expected to answer *yes*.

(3) You know how it is there early in the morning in Havana with the bums still asleep against the walls of the buildings; before even the ice wagons come by with ice for the bars? (1)

The final sentence of the novel likewise uses *you*, this time embedded in description and weighted towards the generic interpretation.

(4) A large white yacht was coming into the harbor and seven miles out on the horizon you could see a tanker, small and neat in profile against the blue sea, hugging the reef as she made to the westward to keep from wasting fuel against the stream. (261)

The framing *you* of J. D. Salinger's *The catcher in the rye* (1951) directly involves the reader in the speech act that constitutes the novel—Holden Caulfield's telling of his story.

(5) If you really want to hear about it, the first thing you'll probably want to know is where I was born, and what my lousy childhood was like, and how my parents were occupied and all before they had me, and all that David Copperfield kind of crap, but I don't feel like going into it, if you want to know the truth. (1)

In the final two sentences, the act of telling takes on universal dimensions with the use of an imperative construction (which implies a *you* subject) and then generic *you.*

(6) Don't ever tell anybody anything. If you do, you start missing everybody. (214)

In the two novels just cited, *you* seems to slide between 'addressee' and 'anyone'. In other works, *you* seems not to have an addressing function but a generalizing and at the same time personalizing one. An example is seen in (7), from J. K. Rowling's *Harry Potter and the sorcerer's stone* (1998).

(7) Mr. Dursley couldn't bear people who dressed in funny clothes—the getups you saw on young people! (3)

The involvement of the reader that is achieved by the use of *you* in fiction can be viewed as part of the attempt of authors to incorporate and represent colloquial language in writing, to make the relationship between writer and reader more akin to a personal, face-to-face relationship between speaker and hearer.

Instances of *you* without a specific addressee are frequent and hardly noticeable in ordinary oral language use, but until recently it was judged a stylistic flaw for written English to include colloquial features, except for fiction and very informal genres. Lately, however, writers of good, clear prose about serious topics use *you*.

From Thomas Friedman's *From Beirut to Jerusalem* (1989), winner of the National Book Award, comes (8):

(8) They say it was a cold, drizzly winter night, the kind that you often find in Beirut or Jerusalem, where the combination of wind and rain leaves you chilled to the bone and wishing you never had stepped outside. (81)

The opening page of Robin Lakoff's *Language Wars*, published by a prestigious university press, contains (9):

(9) The popular use of "linguist" is very different from its professional acceptation. If you tell a layperson that you're a "linguist," you are very likely to be asked, "And how many languages do you speak?" (1)

Feature articles in newspapers often take the *you* perspective, as in 10 from a feature about pregnancy and in 11 about the winner of the university's distinguished service award.

(10) After years of trying, you finally conceive and all you hear from people is, "It's about time." ([New Orleans] *Times-Picayune,* 31 July 2000, C1)

(11) So why quit work you still love? Why say goodbye to people you will miss? ([University of North Carolina at Chapel Hill] *University Gazette*, 16 August 2000, 7)

You seldom occurs in news articles or editorials, except in direct quotation. If the speech reported between quotation marks in newspapers and magazines is a reliable record, many speakers use *you* where the actual referent is the speaker, *I*.

(12) Referring to Smith's 1997 retirement after 36 years, Guthridge said, "I know how Dean felt three years ago. You just are exhausted." ([Raleigh] *News and Observer*, 1 July 2000, A19)
(13) [Quoting Jennifer Aniston] "It's scary how Hollywood treats you like this completely different person when you're thin." (*McCall's*, June 2000, 25)
(14) It has been an exhausting stretch for Doherty. "At night, you collapse." ([Raleigh] *News and Observer*, 19 July 2000, C1)
(15) Why work until you're 83? For Felton, the answer was simple. "It was just fun," she said. "And besides, what else can you do?" ([University of North Carolina at Chapel Hill] *University Gazette,* 16 August 2000, 1)
(16) "You have a routine when you first hit the beach," he said. "You cast a couple of times off the beach. If you get a few strikes, go in ankle deep. But if I don't have to go any deeper, that's fine with me." ([New Orleans] *Times-Picayune*, 2 August 2000, D4)

Other instances of *you* in direct quotation in newspapers are the generalizing type, referring to 'any reasonable person in the situation'.

(17) "When you run, you feel better about yourself," Galloway said. ([New Orleans] *Times-Picayune*, 2 August 2000, D5)
(18) There may be "one-of-a-kind artifacts" still undiscovered, [Morgan] Perkins said, "You'll find gold chains that no one else has seen." ([Raleigh] *News and Observer*, 20 July 2000, B8)
(19) "You download music when you just like that one song off a CD," said Marla [Piece], 15. "You're not going to buy a whole CD if you only like one song." ([New Orleans] *Times-Picayune,* 28 July 2000, A10)

Uses of *you* like those in (12)–(19) are frequent and ordinary in spoken English. The contexts in which indefinite *you* occurs are often not highly informal, and the use of *you* does not make them so. *You* is much more than an informal variant of *one*. It allows a speaker to use a personal form to imply that an agent is like anyone, including the hearer. Rather than occurring in contexts of trivial conversation, *you* is often the choice in the consultative register when the speaker is trying to make the hearer understand something unfamiliar, new, or complex. Examples (20)–(24) were uttered by public figures within the structured but unscripted format of public affairs television broadcasts.

(20) Just because someone would break the law doesn't mean that you shouldn't make the law. (Steve Forbes, *Meet the Press,* 23 January 2000)
(21) It fundamentally changes the way you manage a conflict. (Colin Powell, *CNN 20 Live*, 2 June 2000)
(22) You need to have a chain of command. (Leon Panetta, 'Role of the White House Chief of Staff', *C-Span*, 15 June 2000)

(23) We [the Supreme Court] say you can't have an undue burden. (Linda Greenhouse, *Washington Week in Review*, 30 June 2000)

(24) [The Russian bear might not be out there in the woods any more, but the Chinese dragon is. (Steve Roberts)] How do you run against the Chinese dragon? (John Harris, *Washington Week in Review,* 7 May 1999)

The Michigan Corpus of Academic Spoken English [MICASE] (Simpson et al. 1999) is a particularly rich source for *you* in situations in which the transmission of information is the goal of the communication. MICASE makes available on the web written transcripts of a variety of academic events at the University of Michigan; the current corpus contains twenty-three events, totaling 222,100 words. In the MICASE transcripts are hundreds of occurrences of *you* not directed to an addressee. Examples (25)–(28) are drawn from different events.

(25) The history of science has always been that you have to be first, you have to make that discovery first, and that sets up an incredibly competitive climate, and atmosphere. (MICASE, Women in Science Panel, 17 October 1998)

(26) You have to be able to put names on the organisms. (MICASE, Interview with Mycologist, 4 January 1996)

(27) You [pregnant Inuits] have to be airlifted to the hospital. (MICASE, Medical Anthropology Lecture, 25 February 1998)

(28) On the other hand, if you're on an island where your primary diet is going to be these little insects that bore in, isn't it great that you have this long thin beak? (MICASE, Psychology 111 Lecture, 22 September 1998) [Obviously it is a bird that has the beak, not the students listening to the lecture.]

The tendency of *you* to widen to refer to 'everyone' is shown by the number of proverb-like statements that use *you*, as in 29 and 30.

(29) You can lead a horse to water but you can't make him drink.

(30) You gotta know when to hold 'em and know when to fold 'em.

In the consultative register, *you* is often combined with existential *have* or *got*, as in (31)–(33). In these instances the reference of *you* has moved as far as possible from the particular addressee end of the continuum to generalize as 'existence' or 'topic'.

(31) You have open borders in Europe now. (Georgie Anne Geyer, *Washington Week in Review,* 11 February 2000)

(32) …in a lake such as Douglas Lake that stratifies and you have a thermocline. (MICASE, Biology and Ecology of Fishes Lecture, 21 July 1998)
(33) You've got three million Palestinian refugees. (Martha Raddatz, *Washington Week in Review*, 21 July 2000)

Sentences (31)–(33) can be paraphrased with *there is/ there are* in which another deictic, *there,* plays an equivalent grammatical and semantic role to *you*, (34)–(36).

(34) There are open borders in Europe now.
(35) …in lakes such as Douglas Lake that stratifies and there is a thermocline.
(36) There are three million Palestinian refugees.

The genitive form *your*, as in (37)–(38) from public affairs television programming, requires scrutiny beyond the space limitations of this paper. Because Wales discusses *your* in detail and traces its semantic development to 'typical member of a class'(15), I shall not treat *your* further.

(37) These guys are not your classic teamster thugs. (Glen Simpson, *Washington Week in Review,* 19 September 1997)
(38) Many of your northern species—they travel as far south as North Carolina. (Betsy Bennett, *North Carolina People*, 31 March 2000)

In summary, in contemporary American English indefinite *you* is frequent and unobtrusive. Although it occurs more often in spoken English than in written English, its use does not mark an utterance or a text as highly informal, casual, or insignificant. On the contrary, the indefiniteness of reference actually favors the use of *you* in statements of principles of general applicability. In many contexts, despite indefinite reference, the traces of the personal in *you* help to create an immediacy between speaker and hearer that promotes communication.

As a postscript, I acknowledge choosing the title of this essay more for wit and whimsy than for scholarly rectitude. This study does not provide sufficient evidence for using the adjective *fashionable* to describe the rise or current widespread use of indefinite *you.* How, why, when, where, and to what degree indefinite *you* is increasing and whether or not fashion is a factor are matters for further study, which I intend to fashion.

---

[1] In order to avoid an unnecessarily cumbersome reference section, I have placed bibliographic information needed to locate the numbered examples in parentheses at the end of each example.

[2] I am grateful to my colleague and office neighbor William R. Harmon, author of Macmillan's *A handbook to literature*, for the examples from *To have and have not* and *The catcher in the rye*. I also appreciate the observations of an astute anonymous reviewer that led me to rewrite the final paragraph of the essay.

REFERENCES

American Heritage dictionary of the English language, 3rd ed. 1992. Boston: Houghton-Mifflin.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Bolinger, Dwight. 1979. To catch a metaphor: *you* as norm. *American speech* 54: 194–209.

Garner, Bryan A. 2000. *The Oxford dictionary of American usage and style*. New York: Oxford University Press.

Hemingway, Ernest. 1937. *To have and have not*. New York: Charles Scribner's Sons.

Hodges, John C. et al. (ed.). 1998. *The Harbrace college handbook*, rev. 13th ed. Fort Worth, tx: Harcourt Brace & Co.

Merriam-Webster's collegiate dictionary, 10th ed. 1993. Springfield, ma: Merriam-Webster.

Merriam-Webster's dictionary of English usage. 1989. Springfield, ma: Merriam-Webster, Inc.

Oxford English dictionary, 2nd ed. 1989. Oxford: Clarendon.

Rosen, Leonard J. & Laurence Behrens (ed.). 1992. *The Allyn and Bacon handbook,* 3rd ed. Boston: Allyn and Bacon.

Salinger, J. D. 1951. *The catcher in the rye.* New York: Little Brown.

Simpson, R.C., S. L. Briggs, J. Ovens & J. M. Swales. 1999. *The Michigan corpus of academic spoken English*. Ann Arbor, mi: The Regents of the University of Michigan. [http://www.hti.umich.edu]

Wales, Katie. 1996. *Personal pronouns in present-day English.* Cambridge: Cambridge University Press.

Webster's third new international dictionary of the English language. 1966 [1961]. Springfield, ma: Merriam-Webster.

Whitley, M. Stanley. 1978. Person and number in the use of *we*, *you*, and *they*. *American speech* 53:18–39.

~

WHEN DO SCIENCES MAKE NOUNS FROM VERBS, AND WHEN DON'T THEY?

Carolyn G. Hartnett
Professor Emeritus, College of the Mainland

NOMINALIZATIONS PRESENT MANY PARADOXES: By Zipf's law, longer words should be rare, but in the Longman-Lancaster academic corpus, 4.4% of the words are nominalizations (Biber et al 1998:60). Teachers teach against them, but use them extensively in their own publications (Hartnett 1998). They abound in bureaucracies, which are now trying to avoid them. They played a big role in developing sciences, but new terms in particle physics are not nominalizations. Their use may be declining today.

Systemic Functional linguists call nominalizations grammatical metaphor (abbreviated to *G.M.*), Whorf called them by the limited term *stativations,* and some linguists call them *non-zero-derived deverbal nouns*. I will use the term *nominalizations* now to refer only to nouns made from verbs by changing the root of a verb (usually an old irregular one, such as *sing/song*) or by adding a suffix other than the gerundive *-ing,* which all verbs have. The term *nominalization* exemplifies the process.

To clarify the influences on nominalization, I begin with their patterns of use in a language much different from English, Thai. Then I'll move historically from Greek through the Enlightenment to industrialization and its repercussions and then to patterns of nominalization in current publications. I will discuss written styles in two advancing modern sciences: neuroscience, which uses traditional heavy nominalization for certain audiences, and quantum mechanics, which seems to be moving in the opposite direction. This evidence supports the ten factors that I list.

1. NOMINALIZATION INCREASES WHEN CHANGING TIMES CAUSE INCREASED DISCUSSION OF THEIR TOPICS; IT IS NOT NECESSARILY A CHARACTERISTIC OF THE TOPICS THEMSELVES. In a study of the 700 years of Standard Thai, Amara Prasithrathsint (1997, 1994) found nominalizations since its beginnings, carved in stone in 1283. She saw that nominalization, usually with prefixes, increased during periods of rapid change, such as during the reign of King Rama V (1868–1910) when modernization abolished slavery and reformed education and administration. When things settled down, nominalization decreased. It increased again during the political revolution in the 1930s, when Thailand became a democracy, and later it decreased. Nominalizations were not used equally in all fields: psychology used most (23%), then sociology (18%), science and philosophy (10% each), statistics (8%), mathematics (7%), and medical science (only 5%). She concluded that Thai temporarily increased its use of nominalization not because of the topics themselves but because of increased discussion of them. This is the first factor: language changes in times of change.

2. ONE REASON FOR CREATING NOMINALIZATIONS SEEMS TO BE SIMPLE NECESSITY, FOR WRITERS WANT TO DISCUSS MORE THAN PEOPLE, PLACES, AND THINGS. A spurt in science and philosophy occurred when ancient Greeks incorporated ideas from civilizations to the south and east. They adapted their verbs by adding morphemes that make them mean the actor, the resulting deed or thing done, or the action (Halliday 1998:197). For example, to theorize about the motion of the earth and stars, Greek science created from *move* its words for *mover*, then *moving*, *movement*, and finally *motion,* which is not an actual entity but is treated as stable, measurable, and able to be modified and participate in further processes. Greek science and mathematics needed ordered sets of abstract technical terms like these; the need for new terms is the second factor.

3. NOMINALIZATION IS EFFICIENT, ESPECIALLY WHEN TIME, ACTOR, AND PROBABILITY ARE IRRELEVANT. Western science changed little until the period of the Enlightenment, when Newton challenged the prevailing attitude and revolutionized science by beginning to base conclusions on observations. In 1704, he used a verb, *refracted,* first and then a nominalization of it, *Refraction*.

(1) Light will not be refracted enough, and for want of a sufficient Refraction will not converge to the bottom of the Eye. (Halliday 1998:201)

This sort of nominalization is easy to understand because it grows out of the text, a process that Halliday calls *logogenesis*. It is a natural progression from reporting a factual observation of an event to naming the process. A third reason for using nominalization is its efficiency.

4. NOMINALIZATIONS ALLOW KNOWLEDGE TO ACCUMULATE BY BUILDING ON PREVIOUS INFORMATION, COMPRESSED, ORGANIZED, AND PACKAGED AS A BASE FOR DEVELOPING NEW KNOWLEDGE. In the introduction to his ground-breaking work, *Elements of Chemistry*, Lavoisier recognized 'the impossibility of separating the nomenclature of a science from the science itself: …as ideas are preserved and communicated by means of words, it necessarily follows that we cannot improve the language of any science without at the same time improving the science itself; neither can we, on the other hand, improve a science, without improving the language or nomenclature which belongs to it' (1790:xiv–xv). Then when the life sciences developed in the 19th century, biologists also needed to define and exemplify as they constructed new taxonomies. Nominalizations helped them classify their observations of flora and fauna. Chemistry and biology teachers today continue to emphasize the vocabulary of taxonomies because they continue to believe, 'Whoever controls the language controls the science'.

Later, as the social sciences developed, they also needed nominalizations to interpret changes in society. Nominalizations organize known information, freezing it for interpretation in theories or hierarchies (Halliday & Martin 1993:199-233). They yield a noun phrase that can be used in a clause to express a great deal more than is explicit. Thus they allow a smooth flow of new information when it is re-used as given. The fourth factor is the advantage of using previous information as a base for new knowledge.

5. NOMINALIZATIONS CAN ORGANIZE AND SUMMARIZE INFORMATION RHETORICALLY IN WRITTEN COMPARISONS, SEQUENCES, CLASSIFICATIONS, AND MANY TYPES OF GENERALIZATIONS. Nominalization is most frequent in writing, less common in standard speech, and rarest in non-standard varieties. Conversation occurs in real time, usually face-to-face, and without much time for reflection or for collecting separate ideas into generalities or classification systems. In contrast, much writing deals with different topics and is slowly and deliberately created with a different logical structure. Nominalization can reorganize expository writing rhetorically, instead of in the real-world's time and location sequence.

History and literature need nominalizations too, to construct their abstractions and generalizations. Although literature needs verbs to get readers interested in the story of what happened, later, after thoughtful consideration, conclusions need efficient nominalizations to foreground results. In some situations, separate clauses without nominalization are essential to provide poignancy, but in others, nominalizations are required, for full clauses would impede the flow of information (Longacre 1993: 327). Because nominalizations economically generalize in evaluations, introductions, summaries, and conclusions, they appear in many areas beyond academic fields. They serve in law, administration, and bureaucracies because of the fifth factor, the ability of nominalizations to organize information rhetorically.

6. NOMINALIZATION ALLOWS ITS USERS TO BE VAGUE WHEN THAT IS ADVANTAGEOUS TO THEM, BUT VAGUENESS DISADVANTAGES SOME AUDIENCES, AND ITS OVERUSE HAS LED TO A BACKLASH. As technology and industry developed, so did management and governmental bureaucracies, and they adopted the language of science, gaining its prestige. That prestige led bureaucrats to misuse and overuse nominalizations, either from habit or carelessness, to make an impression, or to manipulate readers. Some lawmakers might have felt that they were safer relying on familiar, traditional wordings. As bureaucrats wrote more, they began to talk in their written style. Now, when scientists and managers have become so collaborative that they do more talking with colleagues than writing, their written and spoken styles are merging, and nominalization occurs in both.

A nominal style can limit the amount of information revealed. It allows speakers to generalize without stating or knowing specifics. It also can de-emphasize or hide relevant information. Moreover, although nominalizations turn processes into objects, nothing requires treating them objectively. Nominalizations are generalities, and it has been said that all generalities are false, including this one. The vagueness of nominalizations is such a great disadvantage to the audience that a backlash eventually developed, opposing the style called *gobbledygook*. Now our government has issued guidelines and is making an effort toward plain English in documents for the public. Many publications, such as insurance policies and health guidelines, have become easier for consumers to understand; improvement is possible. The sixth factor is the vagueness of nominalizations, which can benefit the users but harm their audiences.

7. BECAUSE USE OF NOMINALIZATIONS REVEALS AN ASSUMPTION THAT THE AUDIENCE UNDERSTANDS THEM, THEY INDICATE GROUP MEMBERSHIP THAT INCLUDES SPEAKER AND AUDIENCE. Although technology and business can seem impersonal, nominalization convinces readers of their understanding and thus establishes the cohesion of group identity. Insiders need their own efficient jargon and standard technical terminology, and informed audiences dislike being talked down to. Technical jargon creates a field. Because nominalization requires a knowledge of the field, it distinguishes the expert from the uninitiated. It indicates status, authority, maturity, and other aspects of the writer's identity (Ivanič 1998:135, Hartnett 1999a). Heavy nominalization makes a text sound formal and impersonal, often with a learned touch. It can be interpreted as revealing the writer's assumption that readers understand the content, implying group identity and perhaps boosting the egos of the readers. The seventh factor is that nominalizations assume understanding and therefore reflect the group membership of audience and speaker.

8. ALTHOUGH THE MEANING OF NOMINALIZATIONS IS INHERENTLY DIFFICULT, IT MUST BE MADE CLEAR TO AUDIENCES WITH DIFFERENT NEEDS. Because nominalizations assume previous knowledge, they can increase the difficulty of comprehension. They leave implicit the experiential meanings that beginning learners most depend on (Halliday & Martin 1993:41, 250–57). Nominalizations presuppose so much prior knowledge that they are difficult for students to comprehend; this may be why many educators dislike them. But unpacking nominalization makes texts longer, although still not clear, because it gives all details equal importance and may still hide relationships. Brevity is both advantageous and problematic.

Context and modifiers may be insufficient to clarify what the nominalizations refer to: the actor, action, recipient, resulting product, or more abstractly, the outcome or a quality. For example, the words *corrosion* and *classification* can mean either processes or their products, but *information* means only the product, not the process. *Animal protection* can mean using guard dogs or saving whales. Meanings can change unpredictably. The suffix *-er (player, speaker*), for example, usually refers to an actor or agent, but *laughter* and *prayer* are products, not actors. Even a native reader cannot deduce meaning simply from the form. A *conception* can be an idea, its beginning, a biological process, or its result.

Furthermore, the original speakers and writers have their own intended meanings before they express them, and so the meanings of their nominalizations are 'never simply additive', but are more specific and applicable to a narrower field (Bolinger 1968:55). When scientists try to avoid creating too many new words, they may slant,

Publication, Date (1998), Price, and Number of Articles	**Sentences in First Paragraphs**	**Total of Finite Verbs**	**Total Nominalizations**	**Nominalizations as % of Verbs**
Science: May 8, $7, 18				
• Specialist Reports	117	193	178	92%
• Non-Specialist Reports	64	125	55	44%
Modern Maturity, Mar–Apr, $.40, 19	68	125	57	46%
Popular Mechanics, June, $2.95, 18	79	122	52	43%
PC World, June, $2.95, 18	37	64	27	42%
Family PC, June, $1.99, 9	28	66	28	42%
Essence, June, $2.75, 8	100	89	28	31%

	Specialized	**Non-Specialized**
Total Words Analyzed	687	681
Sentences	35	31
Complete Finite Verbs	67	80
All Nouns	184	174
Nominalizations	41	25
Prepositional Phrases	84	66
• Objects of Prepositions that are Nominalizations	18	3
• Sequences of 2 or more PPs	19	11
• Nominalizations in Sequences of 2 or more PPs	7	1

Table 2. *Analysis of two science reports of the same research.*

science magazines (*PC World, Family PC,* and *Popular Science*), near the 46% in *Modern Maturity* for readers over age fifty but way above the 31% in the extensively illustrated beauty magazine, *Essence*.

When scientists write for other scientists, they usually try to be convincing by staying within the paradigm, but science writing for the public focuses on novelty (Fahnestock 1986). Therefore I examined reports for two different audiences concerning the same brain surgery (Baynes et al. and Strauss, both in *Science* 8 May 1998) to compare the styles and contents appropriate for different audiences and purposes.

To make the specialist report as similar as possible in content and size to the article for educated non-specialist readers, I omitted much of the lengthy narrative of procedures and test results from the middle of the specialist report (Hartnett 1999b, 1999c). Table 2 reflects the more nominal style of the article for specialists. Neither article had many nominalizations that are technical terms (such as *lateralization*, 10 total), for most of the technical terms were adjectives, verbs, or ordinary nouns used in a special sense, such as *hemisphere.* Most of the nominalizations in this very limited sample were clear in context, and only one related to previous use as a verb.

Although Table 2 shows that the specialized material has only 5% more nouns (184 vs. 174), more of them are nominalizations (41 vs. 25, which equals 22% vs. 14% of all the nouns and amounts to 1.2 vs. 0.8 nouns per sentence and twice as many nominalizations per verb, .62 vs. .31).

Differences in the syntactic roles of nominalizations also are significant ($p = .01$), for the report for specialists has more prepositional phrases (84 vs. 66), and six times as many nominalizations in prepositional phrases as the parallel report for non-specialists (18 vs. 3). In the specialist report more prepositional phrases occur in sequences of two or more: the 19 sequences there include seven nominalizations, while the 11 sequences in the non-specialized material used only one nominalization.

Prepositional phrases pack more lexical content into clauses. This heavy modification also simplifies the syntactic complexity without reducing the lexical density (the ratio of content words to grammatical terms). In spoken English, lexical density is only about two lexical words per clause (Halliday 1998: 207). In contrast, both of the science articles that I examined have very high lexical densities of 10.8 content words per average sentence of 20.7 words, for they both are written for scientifically trained readers. In both, 52% of their words serve as nouns, verbs, adjectives, or adverbs, and 48% work as prepositions, conjunctions, articles, or modal auxiliaries of verbs. Heavily modified nominalizations appear in simple sentence structures for the specialists who know the vocabulary, while other readers get fewer nominalizations but more complex syntax. These differences illustrate the eighth factor, that lexical and syntactic choices vary to accommodate the needs of the readers as well as the purposes of the writer.

9. BECAUSE NEW INFORMATION IS EASIER TO COMPREHEND AT THE END OF SENTENCES, NOMINALIZATIONS THAT PRESENT NEW INFORMATION AT THE BEGINNING OF SENTENCES ARE DIFFICULT TO COMPREHEND AND IMPOSSIBLE TO QUESTION. Syntactic roles influence where information appears in the sentence. New information is easier to comprehend after its context introduces it, so it reads better not as subject or theme but in the new information position toward the end. The beginnings of the sentences in the two *Science* reports differed significantly, as illustrated in the structures of the fourth sentences of each report.

(2) By studying an epileptic patient whose brain was surgically divided to control her seizures, Kathleen Baynes, a cognitive neuroscientist at the University of California, Davis, and her colleagues found that the centers for speech and writing, long thought to be in the same side of the brain, can reside in different hemispheres.

(3) Examination of split-brain patients both before and after their surgery is another tool for the unmasking of such processes. [*Author's note: the processes were described two sentences earlier.*]

In the report for non-specialists, the fourth sentence, (2), has fifty-two words; it begins with fourteen words describing the method with a prepositional phrase whose object is described in a passive clause; then it names the researcher as grammatical subject. After identifying her, her institution, and her co-researchers, it continues with the main verb followed by a clause whose subject is followed by a prepositional phrase and a participial modifier with two more prepositional phrases before the verb and another prepositional phrase.

In the specialized article, the nineteen-word fourth sentence, (3), introduces the findings with a nominalization as subject. This sentence exemplifies a sentence structure that is common in science writing: two heavily modified noun phrases, often nominalizations, connected with a verb that relates them. If the specialized example sentence were stripped of the fifteen words in its four prepositional phrases, it would read simply and insignificantly.

(4) Examination… is another tool…

In its entirety, this sentence illustrates how nominalizations attract modifiers of all types, especially noun adjuncts and prepositional phrases, in a relatively simple sentence structure. In contrast, the nominalizations for non-specialists in (2) are the unmodified common words *seizures* and *speech;* both are used as objects, not subjects. These sentences reflect the importance of syntax and word order in the ninth factor, balancing the difficulties of comprehending nominalizations that present new information at the beginning of a sentence.

10. ALTHOUGH WE ARE ACCUSTOMED TO INTERPRETING THE UNIVERSE AS A PLACE WHERE THINGS RELATE TO OTHER THINGS, AND NOMINALIZATION INTERPRETS OR RECONSTRUCTS IT AS SUCH, WHEN OUR VIEW CHANGES, LANGUAGE MUST CHANGE, BUT NOMINALIZATIONS ARE NOT THE ONLY ALTERNATIVE. Nominalizations gave science the advantage of discussing processes and relationships as things that are treated as stable objects and as the basis for further knowledge and theory. Now we have a paradox in the science with the greatest changes in the last century. Although the language of neuroscience and most other sciences reflects a fixed, determinate reality of objects to be observed and classified, physics moved 'from object to process, from absolute to relative, from determinate to probabilistic, from stability to flow' (Halliday & Martin 1993:20). Nobel Prize-winner Steven Weinberg writes, 'quantum mechanics has transformed the very language we use to describe nature: in place of particles with definite positions and velocities, we have learned to speak of wave functions and probabilities… matter has lost its central role' (1992:3). How can physicists use nouns when 'matter has lost its central role'? They describe basic particles as vibrating strings of positive and negative fields of energy. They accept the principle that their knowledge is always incomplete, which follows from the Heisenberg principle of not being able to observe something without changing it (such as the speed and location of a particle).

Quantum physicists see order in the universe but accept flux and use processes that they call *fuzzy logic* and *chaos theory*. Instead of creating new nominalizations for processes or their products, they use humor to name the characteristics of particles. They discuss *anti-bottom quarks* with *flavors* of *charm* and *strangeness*. Some of the whimsical new terms are short combinations of morphemes, such as *gluons* that hold quarks together. *Quantum chromodynamics* classifies with the names of nonexistent colors (Gribbin 1998). These are the formal terms used in specialists' reports, terms defined seriously in dictionaries and labeled not *colloquial* but *physics.* Elsewhere, imprecision is becoming acceptable. Statistical sampling is considered more accurate than an actual census count. A tendency toward imprecision might be a temporary stage for teenagers who go 'like uhm, y'know, well, anyway, like, so, y'know, whatever. It's awesome'.

Our grammar construes a universe of things and relations, where objects relate to or act on other objects. Nominalization interprets the world this way (Halliday &

Martin 1993:220). Our grammar imposes grammatical categories, so when the things are events, we adjust by turning verbs into nouns. However, our view can change and make us change the categorization. The tenth factor is our changing view of the universe, which in turn changes our language.

Another change seems to be developing, making science more commonplace and democratic. Scientists are human beings who need public support for funding and who enjoy having non-scientist friends understand or at least appreciate their work. Halliday has predicted a return to the more democratic use of verbs, and he has also reminded us of the varying categories in non-Indo-European languages (1998:186–88). Language is changing, but not necessarily increasing nominalization. Linguists cannot tell an advancing science what kind of language to use, but they can observe it.

REFERENCES

Baynes, Kathleen et al. 1998. Modular organization of cognitive systems masked by interhemispheric integration. *Science* 280:902–05.

Biber, Douglas, Susan Conrad & Randi Rippen. 1998. *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Bolinger, Dwight. 1968. *Aspects of language*. New York: Harcourt, Brace & World.

Fahnestock, Jeanne. 1986. Accommodating science: The rhetorical life of scientific facts. *Written communication* 3/3:275–96 [reprinted 1998, 15/3:330–50.]

Gribbin, John. 1998. Q *is for quantum: Particle physics from A to Z*. London: Weidenfeld & Nicolson.

Halliday, M. A. K. 1998. Things and relations. In *Reading science*, ed. by J. R. Martin & Robert Veel, 185–235. London: Routledge.

——— & J. R. Martin. 1993. *Writing science: Literacy and discursive power.* London: Falmer.

Hartnett, Carolyn G. 1998. The English nominalization paradox. Presentation at the Linguistic Association of the Southwest. Tempe, az. [eric Clearinghouse on Reading and Communication Skills ed #426411]

———. 1999a. Nominalization: Hiding the actor while making the writer and the reader visible. Presentation at the Conference on College Composition and Communication. Atlanta ga.

———. 1999b. What should we teach about the paradoxes of English nominalization? Presentation at International Systemic Functional Congress. Singapore. In press.

———. 1999c. Twenty-five centuries of nouning verbs: Past, present, and future. Presentation at Linguistics Association of the Southwest, San Antonio tx.

Ivanič, Roz. 1998. *Writing and identity: The discoursal construction of identity in academic writing*. Philadelphia: Benjamins.

Lavoisier, Antoine-Laurent. 1790. *Elements of chemistry*, trans. Robert Kerr. Edinburgh: William Creech. [reprint 1965, New York: Dover.]

PRASITHRATHSINT, AMARA. 1994. Borrowing and nominalization of technical terms in Standard Thai. In *Language reform: History and future*, ed. by István Fodor & Claude Hagège, 9–24. Hamburg: Helmut Buske Verlag.

———. 1997. The emergence and development of abstract nominalization in Standard Thai. In *Southeast Asian linguistic studies in honor of Vichin Panupong*, ed. by Arthur S. Abramson, 179–90. Bangkok: Chulalonghorn University Press.

STRAUSS, EVELYN. 1998. Writing, speech separated in split brain. *Science* 289:827.

WEINBERG, STEVEN. 1992. *Dreams of a final theory.* New York: Pantheon.

WHORF, BENJAMIN LEE. 1956. *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*, ed. by John B. Carroll. Cambridge: MIT Press.

◆

LEXICALIZATION AND OPACITY

Aya Katz
Providence University

introduction. Lexicalization is a process whereby new lexemes are formed. The process is usually seen as historical, and the resultant lexeme becomes a synchronic unit. A certain degree of componential opacity is associated with a fully lexicalized item. (Heine et al. 1991:95) To the extent that speakers still view a lexeme as composed of its parts, the lexicalization process may be considered less than complete within a functional theoretical framework. Such an analysis is consonant with certain typologies, but flies in the face of synchronic evidence in languages where componential transparency is the norm rather than the exception.

For example, it is well known that Chinese has compound lexemes (Li & Thompson 1998). However, there is evidence from brain damaged patients that a component's status as verbal or nominal can affect access to same within a fully lexicalized compound. So Mandarin 吃飯 *chi fan* 'eat' , while enjoying full lexeme status, is more likely to be retrieved as just 飯 *fan* 'rice' by patients with damage to Broca's area, and as 吃 *chi* 'eat' alone by those suffering from damage to Wernicke's area (Bates & Chen 1991; Zhang ms.). On the other hand, speakers of English often have difficulty accessing the componential makeup of words whose root morphemes have independent lexemic status: e.g., *rooster* [< *roost* + *er*]. In Hebrew most morphemes are bound, because consonants carry mostly lexico-semantic information, while vowels code mostly grammatical and categorization data. Despite the highly dependent nature of Hebrew morphology, speakers find it extremely transparent.

The purpose of this paper is to determine the implications for lexicalization from the relative opacity or transparency of morphemes in typologically distinct languages. Lexicality will be defined independently from componential opacity and morphological dependence. Opacity will be subcategorized as circumstantial (objective) versus psychological (subjective). Correlations between and among opacity, dependence and lexicality will then be tracked.

This paper will use examples from Hebrew, English and Chinese to determine how overall systematicity in lexical derivation affects speakers' strategies in processing lexemes. The actual transparency/opacity of a particular item may be seen to be less important than whether a speaker expects lexemes to be componentially transparent. We will consider whether systematicity in derivation primes a subject to use componential analysis even where it is less productive. We will explore whether lack of systematic derivation overall deters a speaker from using a strategy of componential analysis even on lexemes whose components are well known to the speaker. The psychological underpinnings of opacity will be explored within a framework of typological comparison.

1. THE THEORETICAL BACKGROUND. *Lexicalization*, as a linguistic label, emerges from the functional linguistic school of thought, and it is often paired with the term *grammaticalization*. (Heine et al. 1991; Hopper & Traugott 1993; Traugott & Heine 1991). As historical processes, both lexicalization and grammaticalization often result from reanalysis. However, while grammaticalization gives rise to new productive patterns that add to the grammar of a language, lexicalization produces seemingly isolated additions to the lexicon. Regularity, productivity, and transparency as components are therefore attributed to grammaticalized items, whereas idiosyncracy, randomness and opacity are expected of lexicalized words.

According to Heine et al. (1991:95), the difference between the two terms can be summed up as follows: 'Assuming that both involve some kind of 'idiomization', the latter may be said to be morphologically productive in the case of grammaticalization but not in that of lexicalization.'

The above is representative of statements about lexicalization in grammaticalization literature as a whole, and it carries with it some subtle implications that are worth exploring. There is also an almost imperceptible shift between a diachronic use of the term to a synchronic use that has far reaching implications. An item is often seen to be more lexicalized if its componential makeup is relatively opaque, and less lexicalized if it is relatively transparent. This implies that full lexemic status ought to be accorded only to items whose derivation is synchronically unreadable. If we followed such logic to its ultimate conclusions, we would decide that a monomorphemic word is the most lexemic of all words. However, there are serious drawbacks to this approach when applied crosslinguistically to languages belonging to contrasting typologies.

In a language such as Hebrew, there are almost no monomorphemic words, since vowels and consonants play complementary roles in word formation. The majority of vowels found in a Hebrew word code derivational and inflectional contrasts, while a majority of consonants code lexico-semantic information. Morphemes are discontinuous, and every syllable carries parts of more than one morpheme. But despite the extremely fusional nature of the language, componential transparency is the norm, rather than the exception.

In fact, in Modern Hebrew, non-linear derivation patterns, involving the use of discontinuous roots and templates to form new words, are still extremely productive. (Bolozky 1999).

Mandarin is known as an isolating language, but it has a large stock of complex lexemes. (Li & Thompson 1998). The dependence level of these lexemes is low, in that they are not phonologically affixed and there is no reduction or resyllabification involved. In this sense, the morphologically complex lexemes of Mandarin are quite different from those of Hebrew. But both Mandarin and Hebrew enjoy a high level of morphological transparency. There is evidence from brain damaged speakers of Mandarin that components are accessed according to their grammatical categorization, rather than the grammatical category of the word as a whole (Bates & Chen 1991).

Multimorphemic words in English are more tightly fused than those in Mandarin, but less so than in Hebrew. Multisyllabic lexemes in English are common, but there is by no means a one-to-one correspondence between morpheme and syllable. Speakers of English have great difficulty analyzing lexemes into their component morphemes. Even in such common words as *heal* and *healthy,* speakers of the language require instruction in its history in order to identify the common root, according to the introduction to a leading textbook on the history of English (Pyles & Algeo 1993.2) In other words, commonly used English lexemes are often componentially opaque for the average speaker.

The implications for the concept of lexicalization from the above observations of Hebrew, Mandarin and English are as follows:

(1) Lexical status is not necessarily dependent on degree of fusion.
(2) Lexical status is not necessarily dependent on opacity or semantic bleaching of components.
(3) Opacity and fusion are independent of each other.

The significance of these initial observations is as follows: if lexical status can be established independently of both fusion and opacity, then lexeme formation (i.e., *lexicalization*) cannot be defined as a necessarily 'morphologically unproductive'.

Heine et al.'s statement (1991:95) might suggest that componential opacity is something that indicates a high level of lexicalization, where the term lexicalization is defined as incremental. The higher the opacity, the higher the degree of lexicalization. But if such an interpretation were used, most words in a language such as English might be judged to be more lexicalized than those in Hebrew and Mandarin. What would be the implications of such a finding?

One possibility would be to suppose the individual words have undergone a greater process of semantic erosion and that opacity results from such bleaching. But an alternative explanation is available: it is not so much that individual words have had their morphemes bleached, but rather that the speakers of the language, due to lack of overall systematicity in lexical patterning, have been conditioned to ignore componential analysis as a useful strategy.

This paper proposes to probe the issue by comparing psychological componential opacity (the psychological ability of speakers to analyze components) with circumstantial componential opacity (whether the requisite cues are synchronically there in the language to allow for such analysis).

2. PSYCHOLOGICAL COMPONENTIAL OPACITY VERSUS CIRCUMSTANTIAL COMPONENTIAL OPACITY. An extreme example of circumstantially opaque derivation in English is *hussy* [<*housewife*]. (Pyles & Algeo 1993:265) The phonological reduction and fusion that occurred over time have rendered the lexeme morphologically opaque

and therefore monomorphemic. Speakers of English cannot identify the component parts of the word, because no one could do so without knowing its history. The information is simply not available synchronically within the language as a whole. Therefore the opacity is circumstantial, not psychological.

An example of psychological opacity in English can be found in the lexeme *rooster* [<*roost* +*er*]. Here, the morpheme *roost* has separate lexemic status, and the suffix *-er* is productive throughout the language. Speakers synchronically have enough information to easily analyze the word. If they fail to recognize its components, this is due to psychological opacity. (That many do fail to recognize the components has been related to me by a number of native speakers of English who first learned of this word's composition in a morphology class.)

The methodology proposed has at its core the comparison of psychological versus circumstantial opacity. If a lexical item is both circumstantially and psychologically opaque, there is nothing surprising in such a situation. Likewise, if a lexeme is both circumstantially and psychologically transparent, there is nothing unusual about that. However, if there are lexemes that are psychologically opaque but circumstantially transparent, such a finding would have theoretical significance. Moreover, if a lexical item is circumstantially opaque but psychologically transparent, this would be a very interesting find.

3. PROPER NAMES AND THEIR IMPLICATIONS FOR LANGUAGE-WIDE COMPONENTIAL TRANSPARENCY. It might be supposed that opacity or transparency of lexemes would be largely determined by the necessity of engaging in creative acts of word formation. Since most speakers seldom have occasion to coin new words, their skills in the reverse process of analyzing the words with which they are presented might be atrophied.

The generative framework, for instance, equates creativity with the ability to generate novel utterances. (Chomsky 1966). The focus is on the speaker, not the hearer. Is transparency directly related to productivity? Some of the literature might suggest this.

Bolozky (1999:xi) tells us in his preface: 'Somewhere in the heart of every speaker of Israeli Hebrew, placed between a love of oriental food and a passion for politics, there is a constant urge to form new words'. He then proceeds to test productivity by a series of experiments in which speakers are requested to coin new words, and the templates and roots which they use are analyzed for productivity.

The approach of this paper is quite the reverse. It is speakers' comprehension of the componential makeup of commonly used lexemes that concerns us here. Nevertheless, the one time when most ordinary people have occasion to form a word is in the process of naming. A proper name is the most extreme form of lexicalization available. Since many speakers do engage in naming, they must also have some skill in deducing the meanings of names given by others. A brief examination is in order, in the spirit of a crosslinguistic comparison.

The following quoted material is of a non-academic nature, and it is presented not in order to prove any linguistic assertions contained therein, but to demonstrate the attitudes toward names displayed by ordinary lay people. In a short article, intended for Chinese speaking students of English, Oscar Chiang explains: 'The Chinese are more particular than most Americans when choosing personal names. Chinese names often have special meanings, while most American names do not. Chinese parents or grandparents usually select names for their children or grandchildren with great care in order to create a certain expectation or ideal for the child to follow. On the other hand, American parents adopt the names of relatives (and even friends) for their children, as a way of honoring that person'. (Chiang 1999).

The facts on which the above lay perception is based are as follows: most names used by English speakers are not English words. Most are borrowings from other languages, and even many whose origin is Anglo-Saxon are not synchronically recognizable English words. As a result, English speakers have been conditioned to think of names as semantically empty labels. The attitude is so systematic for English speakers that it colors their perception even of names that are circumstantially transparent in modern day English.

There is evidence that at least some speakers of English, while being perfectly aware of the meaning of the word *rose,* appear to be concurrently unaware that the name *Rose* also means 'rose'.

In a web search of name dictionaries, one such dictionary listed the name *Raisa* with the definition 'rose'. The name *Rose,* however, was defined as: 'Of the Rose Blossom'. In other words, the English speakers who compiled that name dictionary were not comfortable with having a name mean the same as a common word with which it is identical. At most, for many speakers of English, a name can have an etymology, but not a lexical meaning.

Since 'meaning' cannot be prescribed, and native speakers of a language are the only ones who determine the meaning of words in their language, another way to express this is to say that the name *Rose,* in the minds of at least some speakers of English, has no meaning directly associated with that of the common noun *rose.*

This assertion may seem slightly controversial, since many English speakers who are linguists have etymologically based perceptions of lexical meaning associated with word origin. Many literate speakers of English are aware that names have 'meanings'—without knowing what those meanings are. The meanings form a set of arcane knowledge. English speakers often consult name dictionaries to find out what their names 'mean'. What they are really looking up is etymologies, true or fabricated.

The same desystemetizing influence that we find in the proper noun inventory of English speakers can also be found in the list of common nouns. Borrowed words in English have elaborate morphology which is fairly systematic, but unintelligible to the untrained speaker. Most Germanic common nouns in English are monomorphemic. Barnyard animals make for a good subgroup: *cow, horse, dog, cat, duck, goose, hen, cock.* Words such as *chicken* and *rooster* are in fact bimorphemic—*chick-en, roost-er*. However, due to the lack of systematicity in the language, many speakers do not interpret them as such. A circumstantially transparent word such as *rooster* becomes psychologically opaque to the average English speaker, perhaps because of lack of priming.

4. COMPONENTIAL PRIMING IN CHINESE. In the case of Chinese the morphological components of a multimorphemic lexeme are not fused, by reason of language typology. The phonological independence of each morpheme has led scholars to suppose in the past that each syllable constituted a separate lexeme. It is by now well established that phonological word boundaries and lexemic word boundaries are not necessarily coterminous. (Li & Thompson 1998). Thus, a word such as 吃飯 *chi fan* 'eat' is a single unit for purposes of lexical insertion into a syntactic pattern, despite the phonological independence of its components, 吃 *chi* 'eat' and 飯 *fan* 'rice'. Nonetheless, psycholinguistic evidence shows that the component morphemes continue to enjoy psychological transparency, despite the lexical status of the compound. (Bates & Chen 1991).

This componential make-up of a majority of Chinese lexemes, both nouns and verbs, is systematic in terms of being widespread, but not in terms of generativity. This is derivation, not inflection. We can explain how a word is derived after the fact, but we cannot absolutely predict what derivational paths will be followed. In that sense, there is no anticipatory system for speakers to follow. The patterns are both transparent and productive, but not systematically generative.

Nonetheless, derivation in Chinese is by and large open and accessible to the average speaker. Relative to English, Chinese derivational morphology enjoys a high level of psychological transparency.

This does not, however, mean that every derivation is equally transparent. For instance, the word for China, 中華 *zhong hua,* is composed of two morphemes. Speakers will readily explain that 中 *zhong* means middle, and that China is so called because it is at the center of the world. But the second morpheme 華 *hua* is not as accessible. When questioned about it, speakers look confused and offer a variety of responses, but most admit they are not sure what it means.

As it turns out, the reason for the relative opacity of 華 *hua* directly correlates with word frequency. The word 中 *zhong* is a common word in Modern Mandarin, and hence compound words derived from it are also transparent in terms of the meaning of that component of the compound. The word 華 *hua* is somewhat archaic. It appears in compounds and especially in names, but its more literal meaning is bleached and obscured. Hence, the opacity of the component morpheme is entirely circumstantial. Speakers do not have synchronic access to the word 華 *hua,* and hence they do not know what it contributes semantically to the compound.

The circumstantial opacity described above contrasts nicely with the circumstantial transparency but psychological opacity that English speakers experience with the *roost-* of *rooster.* Chinese speakers do not know what 華 *hua* means, so they also do not know what it contributes to 中華 *zhong hua.* But there are English speakers who do know what *roost* means, yet do not recognize it when it serves as the root for a morphologically complex word.

The non-fusional nature of Chinese phonology may be credited with the ease with which speakers analyze the component morphemes of their multimorphemic words. However, this is not the case in Hebrew, where componential transparency is also the norm, despite the widespread use of discontinuous morphemes.

5. HEBREW MORPHOLOGICAL TRANSPARENCY VS. PHONOLOGICAL OPACITY. In the case of Hebrew, derivational patterns are so predictable and the paradigms are so widespread that one can almost assert that Hebrew derivation is not only systematic, it is downright generative! An average speaker can be given the task of generating a new lexeme using a particular template and root, and in as little time as it takes to conjugate a verb, he can produce a new lexeme. (Bolozky 1999:xi).

Hebrew derivation is not as predictable as Hebrew inflection, of course, but it comes very close. The wide gap between inflection and derivation that European languages typically display is not present in a Semitic language such as Hebrew.

It is well established that Hebrew derivational morphology is extremely productive and transparent to speakers (Bolozky 1999; Ephratt 1997). Therefore, in dealing with Hebrew, this paper will focus not on the question of psychological opacity, but rather on circumstantial opacity. It has been generally assumed that transparency in Hebrew morphology is directly correlated with systematic productivity. The question remains: how does the psychological transparency of a particular lexical item's components survive phonological changes that create circumstantial opacity? Is there circumstantial opacity in current day Israeli Hebrew?

Hebrew phonology has historically had the following stop phonemes /bdg ptk/ which were realized as their corresponding fricatives post-vocalically, and as stops everywhere else. There has been some erosion of this rule since the revival of Hebrew, in that /t/, /d/, and /g/ are never realized as fricatives regardless of environment. Recently there has also been much inconsistency in the realization of /b/, /p/ and /k/ as stops or fricatives. Productively, the stops are not fricativized, but they appear as fricatives in many frozen expressions, including many derived lexemes. This is a situation which ordinarily leads to merger of phonemes and a high level of homonymy among roots. The expected result of such changes is circumstantial opacity of roots.

Thus a lexeme such as /miknasajim/, meaning 'pants', realized as [mixnasajim], may not be circumstantially interpreted as being related to a word such as /knesset/, meaning 'parliament', realized as /kneset/, even though both historically come from the root /k-n-s/, meaning 'to enter'.

The approach to circumstantial opacity adopted here is somewhat akin to the issue of biuniqueness in phonology. Certain phonological theories assume a top down approach, looking at phonemic status from the point of view of the speaker. Requiring biuniqueness in a phonology, on the other hand, means looking at phonemic status from the point of view of the hearer. If the hearer has no way to determine from the sounds whether the [x] of [mixnasajim] is a *het* or a *kaff*, then under a biunique theory of phonology, all the [x] sounds have merged into a single phoneme, and the words derived from the root /k-n-s/ will also be split according to whether the *kaff* in a particular derivation is postvocalic or postconsonantal.

In terms of psychological transparency, such a split has not occurred. Speakers still experience the root /k-n-s/ as an accessible morpheme. They are able to generate new lexemes using this root and familiar templates, even though they may falter in deciding whether to realize the /k/ as a [x] or a [k]. Likewise, when they encounter an established lexeme, the root /k-n-s/ has psychological reality for them, despite the circumstantial opacity created by phonological change. In fact, it is well known that Hebrew displays a concurrent semantic transparency in the face of phonological opacity (Barkai 1978, 1980). Speakers know what a root means and recognize its occurrence, even if they are a little fuzzy as to the phonemes of which it consists. Circumstantial opacity does not prevent them from experiencing psychological transparency.

The apparent implication for lexicalization is this: lexical status is not dependent on componential opacity. In turn, componential opacity or transparency seems to be a systemwide psychological phenomenon. Opacity is not necessarily brought about by the reanalysis of phonological boundaries. Sometimes it is a change in the entire system that renders a particular lexical derivation opaque, despite the circumstantial accessibility of its component parts. Speakers can perhaps tolerate a high level of circumstantial opacity, if derivation in the language as a whole is systematic and predictable.

6. CONCLUSION. Functional research in grammaticalization has equated lexicalization with low productivity, fusion and componential opacity. This paper has explored the possibility, using Chinese, Hebrew and English data, that circumstantial opacity and psychological opacity do not always correlate. We have seen that a lexeme can be psychologically transparent while circumstantially opaque (Hebrew), and conversely, that a lexical item can be circumstantially transparent but psychologically opaque (English). Likewise, Chinese morphological components of compound words were found to behave sometimes as more transparent and sometimes less, without regard to phonological fusion. Thus, the idea that fusion or opacity are predictable correlates of lexemic status has been shown to be questionable.

Lexicalization, as a process, can occur as a result of reanalysis, but reanalysis itself does not necessarily lead to componential opacity. The opacity or transparency of the components of a word may be a psychological effect brought about by language-wide phenomena. Perhaps speakers are conditioned by the language as a whole to use componential analysis as a productive strategy or to fail to make such use.

The above observations are preliminary in nature. I am currently working on an experimental research project funded by the National Science Council of Taiwan. My investigation of componential transparency and opacity is ongoing. Analysis of the data I am currently collecting may eventually support or disprove the initial hypotheses set forth in this paper.

REFERENCES

BARKAI, M. 1978. Phonological opacity vs. semantic transparency: two cases from Israeli Hebrew. *Lingua* 44:363–78.

———. 1980. Aphasic evidence for lexical and phonological representation. *Afroasiatic linguistics* 7:163–1187.

Bates, Elizabeth & Sylvia Chen. 1991. The noun-verb Problem in Chinese aphasia. *Brain and language* 41:203–33.

Berk, Lynn M. 1999. *English syntax: From word to discourse.* Oxford: Oxford University Press.

Bolozky, Shmuel.1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: E. J. Brill.

Chiang, Oscar. 1999. More about names. *Time for students* 15283:10–11.

Chomsky, Noam. 1966. *Topics in the theory of Generative Grammar*. The Hague: Mouton.

Ephratt, M. 1997. The psychological status of the root in Modern Hebrew. *Folia Linguistica* 31:77–103.

Hopper, Paul & Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd, Ulrike Claudi & Friederike Hunnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Li, Charles & Sandra Thompson. 1998. *Mandarin Chinese: a functional reference grammar*. University of California Press.

Pyles, Thomas & John Algeo. 1993. *The origins and development of the English language*. New York: Harcourt, Brace Jovanovich.

Traugott, Elizabeth & Bernd Heine. 1991. *Approaches to grammaticalization*. Amsterdam: John Benjamins.

Zhang, Xiuhong. *Evidence for a neurological distinction between nouns and verbs.* ms.

◆

AN OPERATOR-BASED INTEGRATION OF COMPREHENSION AND GENERATION

Deryle Lonsdale
BYU Department of Linguistics

This paper[1] discusses the theoretical and practical implications of using an operator-based framework for modeling natural-language comprehension and production processes. In particular, it presents aspects of the NL-Soar cognitive modeling system that relate to these questions. It surveys the cognitive theory and modeling system that frame both types of processing, linguistic aspects of the comprehension system and its relevant operators, and the generation system with its own operators and processing stages. The degree to which these two systems exhibit shared functionality is then explored, and ongoing work is mentioned.

1. background. Extensive study has been carried out on the capabilities and shortcomings inherent in human processing of natural language (e.g., Yngve 1960). Psycholinguistic evidence has pointed to observable constraints on complexity and capacity; for example, many types of grammatically correct but unparsable sentences have been described. This has led to discussions about characteristics of the human mind that may play a constraining role in our ability to process language, focusing on such issues as modularity, input/output buffering, working memory size (Gibson 1991; Lewis 1996), parsing error recovery (Pritchett 1992), underlying processes (Gile 1990), and capacity constraints (Just & Carpenter 1992). On the other hand, many computer systems that have been developed to carry out language tasks such as parsing and generation are based on computer architectures that do not always reflect the properties often ascribed to human-based processing of language. Whereas the cognitive plausibility of such systems is not a direct issue in this paper, the underlying assumption is that grounding a generation system in a more general theory of human cognition will lead naturally to an integrated framework where comprehension and generation are both able to draw upon processes, goals, and linguistic knowledge structures that are maximally related. The NL-Soar system, as sketched in this paper, responds to these criteria.

Describing the exercise of human intelligence including (but not limited to) language, deliberation, action, and performance within one general framework, the mind, is the goal of those seeking to establish a unified theory of cognition (Newell 1990; Anderson 1993). Far from being simply a theoretical undertaking, though, this work also seeks to instantiate the approaches as working computational systems in order to demonstrate and explore assumptions about knowledge sources, behavior timelines, resources, and observable mechanisms and processes.

Soar (Newell 1990) is one system that has been used for widely varied cognitive modeling tasks such as game-playing, air-traffic control, and automobile driving. The natural-language processing instantiation of the general Soar architecture is called NL-Soar (Lewis 1993); NL-Soar is thus meant to be the language cognition subcomponent of the general Soar theory. It has been successfully integrated with Soar systems designed to carry out other tasks (Rubinoff & Lehman 1994a; Lehman et al. 1995; VanDyke & Lehman 1997). After a brief introduction to NL-Soar comprehension, the current generation system (which was developed by the author) is sketched, and their integration is then explored.

2. NL-SOAR COMPREHENSION. NL-Soar comprehension was the first comprehensive linguistic component based on the Soar system (Lehman et al. 1991). It directly implements Soar's basic characteristics: a production-rule knowledge base, an agent-oriented infrastructure, a specifiable goal hierarchy, problem spaces that encapsulate and delimit working subsets of memory, and learning via chunked behaviors. It was designed to model the performance of humans while comprehending difficult utterances with multiple embeddings and garden-path sentences; its behavior was shown to closely model that of humans for problematic and unproblematic constructions. More details about Soar and the NL-Soar comprehension system can be found in (Laird et al. 1993) and (Lewis 1993) respectively.

The NL-Soar agent receives textual input sentences, tokenizes them, and stores the resulting words in a temporary buffer that is subject to decay. Words are attended to one-by-one and processed by the system. Processing involves accessing the lexicon for information about each individual word, and then incorporating the word into different types of linguistic structure.

Two different but related linguistic structures are managed by the system. A phrase-structure tree representation called a syntactic model reflects the surface syntactic representation of a sentence. It is based on X-bar syntax and reflects constraints such as those operant in the distribution of topicalized constituents, empty categories, and movement processes involved in inflection. (Chomsky 1986). The semantic representation currently implemented in the system is based on Lexical-Conceptual Structure (LCS) (Jackendoff 1990) and relates actions, states, processes with concepts and their attributes. This structure, called the semantic model, is meant to abstract away from morphosyntactic levels of representation. The use of other semantic representation systems is possible, and in fact others have been used in the past (Newell 1988).

Figure 1 shows syntactic and semantic models that the NL-Soar comprehension component creates upon processing the first part of the input sentence 'The doctor who called works at a hospital'. To begin with, after its introduction to the system the word *the* is recognized as a determiner; after the word *doctor* is processed a NP is formed and the semantic concept *doctor*, a Thing, is spawned. After the next word *who* enters the system, a CP complement is generated for the noun in the syntactic model. After processing the verb *called*, a verb phrase for the embedded clause's subject is created in the syntax, and the appropriate Action node in the semantic model is created and

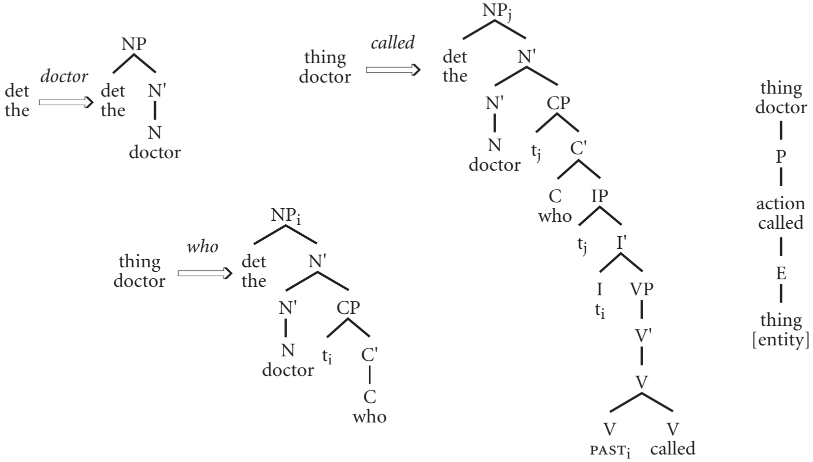


Figure 1. *Syntactic and semantic models incrementally being built after each word in the sentence fragment 'The doctor who called…'.*

attached as a Property to the *doctor* concept. In fact much more information for both models is stored; the figure is only meant to sketch the most superficial attributes of each model in order to give an overall appreciation of the types of structures built.

By design, the system is constrained to only allow one working hypothesis of each kind (syntactic and semantic) as processing progresses. As a result, in certain situations such as garden-path constructions the wrong hypothesis will be pursued; when the current hypothesis becomes untenable in the light of subsequent context, the system is able to perform limited repair by snipping off the erroneous portions of the syntactic model, and then re-building it correctly.

There is a close syntax/semantics correspondence: syntax is prior, and the system's semantic representation is interpreted from it. The category mappings are straightforward: nouns map to basic concepts like Thing, adjectives map to Property, and verbs map to Action, State, or Event. The interaction between syntax and semantics extends to repair: syntactic snips initiate semantic snips which repair the semantic model in light of repairs to the syntax model. Similarly, there are semantics-based motivations for syntactic snips: an untenable semantic model can force previously constructed syntax to be re-analyzed.

All of the above areas of functionality are performed by *operators*[2]. As in many traditional artificial-intelligence-based approaches, NL-Soar uses operators as the instrument for effecting state changes. An operator can be thought as an atomic bundling of actions that reflect some cognitive operation. Comprehension in NL-Soar performs lexical access, parse-tree construction, and semantic network construction piecewise, with (roughly) one operator of each type per word processed. For example, one syntactic operator would be used to combine the determiner *the* and the noun *doctor* into an NP. Similarly, a semantic operator would attach the concept *called* as a Property of the Thing concept *doctor*.

Op.	Operation	Related Constraints	Example
o23	Link a DET as [spec NP]	spec-head agreement, order, unique assignment	the + meeting *a meetings
o47	Link an NP object as [comp V']	subcategorization, order, unique assignment	saw + the cat *sneezed + her
o76	Link an NP subject as [spec IP]	subject-verb agreement, order, unique assignment	dogs + bark *men + sings
o12	Project a node of cat X to XP	V ⇒ V' ⇒ VP	sneezed
o27	Link an NP object as [comp P']	subcategorization, order	to + the store *away + the dog
o39	Link an actor to its action as the external argument	semantic constraints: class, animacy, saturation	Person+Action Thing+Action
o42	Empty operator: note word, record as unintegrable into current context	no actions possible at this stage of processing	the

Table 1*. Sample operator types and the linguistic operations they perform, related constraints, and examples where they would and would not apply.*

Table 1 shows a closer look at the linguistic nature of specific operators. **o23** (the names are mere arbitrary tokens created by the system as they are learned) is an operator that builds the specifier node into a noun-phrase (NP) constituent. In doing so, it must satisfy all of the requisite constraints specified by the grammar of the language; for English, there must be number agreement, the specifier must precede the head noun, and the two pieces must not have been previously assigned to other constituents. This operator would thus fail to build a constituent for the word sequence **a meetings*, but would succeed for *the meeting*. The operator is of sufficient generality as to construct countless other NP's in a similar manner, as long as the linguistic constraints are satisfied. Now consider **o39**, an operator that links together pieces of semantic structure. In a sentence like *the chair sneezed*, one possible meaning interpretation is that a person (e.g., a college department chairman, of semantic class Person) performs the act of sneezing (an Action) and a less plausible one involves a man-made artifact (a chair, of semantic class Thing). This operator tests for semantic class compatibility, in essence testing and enforcing the linguistic constraints often called semantic cooccurrence restrictions. For example, **o39** rules out Thing as the agent of the Action *sneezed*, and instead licenses the semantic attachment with class Person for its agent. Other operators illustrated in the table perform other types of syntactic and semantic structure-building operations. The empty operator (in this case **o42**) notes the presence of a word for which no work can be currently done to integrate it into current syntactic or semantic context, at least until more information is provided (e.g., when new words are input to the system and thus added to the current context).

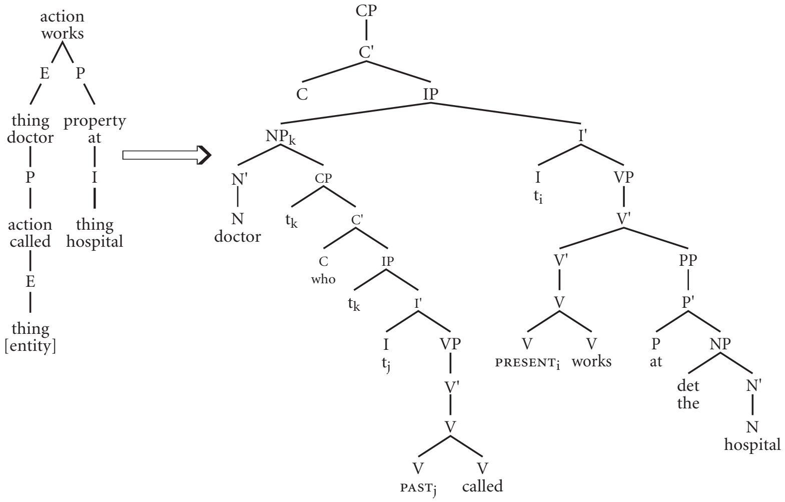


Figure 2. Schematic overview of generation. Operators map pieces of the semantic model at the left into a corresponding syntactic model, depicted at the right.

3. NL-SOAR GENERATION. Given the previous discussion of comprehension in NL-Soar, it will be seen that much of the same principles and methods apply to the NL-Soar generation system. So far, work has focused on generating the output utterance via a syntactic model generated from predefined semantic structure; pre-tactical text planning (i.e. involving high-level considerations such as rhetorical structure, textual cohesion, argumentation strategies, etc.) has not yet been implemented but should be possible given the hierarchical, goal-oriented Soar framework. There are two conditions that initiate tactical generation. The first is that some discourse move, such as responding to a question, prompts the NL-Soar agent to communicate. The second requirement for generation to proceed is that a coherent, connected LCS must exist.

The crux of the current generation process is thus the mapping of semantic structure to a well-formed syntactic utterance model. Once the syntactic structure has been generated, further processing generates the final output word sequence by reading off the lexical items from the terminal nodes of the structure. This results in a word-sequence (i.e. a sentence) which the agent outputs. A schematic representation of the central part of the generation process, mapping from a semantic model to a syntactic model, is given in Figure 2.

Generation of a sentence, then, proceeds as follows. Once the system has determined that a sentence must be generated, it traverses the connected semantic model to identify each concept and relevant annotations (such as tense, aspect, and plurality). All pieces of the model that could potentially contribute lexical material to the eventual output sentence are then considered globally. Particular types of operators, called selection operators, choose and prioritize which pieces of the model to process in which order. English operators, for example, specify that determiners must be generated before nouns, and verbs before their object.

Once a given piece of the semantic model is chosen to process, work can continue in one of two modes. Recognitional processing (i.e., work that has already been 'chunked up', or learned from previous processing) can be carried out via pre-existing generation operators. Generally, one generation operator is used to realize each concept of the semantic model, though the correspondence need not always be strictly one-to-one. When appropriate recognitional operators do not exist *a priori*, more costly deliberate processing must be carried out to create them. This process is sketched in the next paragraph.

Each generation operator can be thought of as a bundled set of actions that allows generation to proceed. Its constituent actions include the following:

1. *lexicalizing the concept:* This stage selects a portion of the semantic model and computes one way of rendering abstract concept(s) into some lexical item(s).
2. *performing lexical access:* The lexicon is then accessed in order to retrieve syn/sem information for the targeted word(s). Words are given categories and projected to form lexical-phrasal items that can then be incorporated into the syntax model.
3. *building syntactic operators:* This step takes the projected word and calculates all possible ways it can be grammatically incorporated into the ongoing syntactic model, keeping track of agreement features, syntactic constraints, constituent structure, and available working memory. Note that any of the sample operators listed in Table 1 for comprehension can also be used in this stage of production. This point will be pursued in the next section.

The three deliberate steps outlined above (lexicalization, access, and syntax model construction) may be bundled into a recognitional generation operator. This allows each of the three stages to fire in association with the others, and in the sequence specified.

Occasionally this three-stage generation operator construction process does not succeed: inappropriate lexicalizations may be chosen, corresponding lexical items may be missing from the lexicon, or incompatibility may arise with the ongoing syntax model. When this happens, generation can recover and re-initiate the process with other selected words. This process continues until a successful outcome is achieved. Thus generation employs a generate-and-test strategy, iterating until a viable generation operator is constructed. This testing is done in lower, temporary problem spaces that are discarded upon completion.

The goal of the NL-Soar generation procedure is to learn operators that render concepts/features from the semantic model into a grammatically correct syntactic model. This approach resembles Dorr's LCS-based approach for machine translation (MT) (Dorr 1993) and subsequent related work in information retrieval; both approaches traverse a conceptual structure to perform incremental surface-language syntactic utterance generation. The operator mechanism is what distinguishes the NL-Soar approach from most other approaches. The lexicalization phase which establishes a correspondence between appropriate pieces of the conceptual model and lexical items was inspired by efforts in realizing pieces of interlingua representation in more conventional, engineering-style knowledge-based system generation components; the lexical realization of objects, of predicates, and of features such as negation, tense, and mood follow principles used in the generation component of an interlingua-based MT system (Lonsdale et al. 1995).

4. INTEGRATING COMPREHENSION AND GENERATION. The version of the generation system described here takes the baseline NL-Soar system as a point of departure. The earliest version of NL-Soar only performed comprehension; later NL-Soar work did include operator-based generation (Rubinoff & Lehman 1994b), but with minimal overlap with the comprehension component. In developing the current generation system, an effort was made to share functionality across these two modalities, thus reducing system complexity. The result of this integration is that many of the same constraints, data structures, and system resources are used for both comprehension and generation. In a very direct and formal way, then, generation is built upon comprehension; the significance of the work reported in this paper is that, as far as this author is aware, no system has yet been developed that exhibits all three of these properties in an architecture that is capable of learning the basic stages of linguistic processing. This section summarizes some areas of shared functionality between the two linguistic processes of comprehension and production:

- *Comprehension and generation both use the same lexicon*. Morphosyntactic and semantic information for each word is encoded in productions that can be used in both processes. Additionally, this allows both to use the same lexical access operator.
- *Comprehension and generation both use the same syntactic operators.* Construction of the syntactic operator, whether during comprehension or generation, proceeds in an identical fashion. Since this integration permits the same type of operators to be used for both, the two processes can bootstrap one another. For example, the operators such as those listed in Table 1 can be made available for processing by both linguistic modalities (comprehension and generation). Thus syntax operators learned in the process of comprehension can be directly used for generation, and vice-versa. This transfer of knowledge between processes opens up many possibilities for exploring facilitation and interference between these two language subtasks. This has important implications for investigation of the human linguistic faculty, both in terms of performance and in competence.
- *Comprehension and generation both use the same architecture.* The decay-susceptible input buffer also stores words for output generation of a sentence. A data structure called the assigners/receivers set, which restricts the number of syntactic constituents that can be stored at a time and computes their compatibility for combination, is shared by both processes. This set

constrains how much linguistic structure can be processed, in cognitively plausible ways (Lewis, 1993).

- *Comprehension and generation both use the same learning mechanisms.* Just as comprehension operators can be learned, so too can generation operators. Furthermore, since operators are compositional (generation operators incorporate lexical access syntax operators), their structure can be built up in turn by incorporating previously learned operators.

Though there are compelling parallels in the architecture and processing methods in the system, this does not preclude the existence of interesting asymmetries. In fact, since some human processing asymmetries have been described in comprehension and generation, we would expect this to be modeled by NL-Soar. Such situations do arise. In comprehension one often encounters certain local, temporary syntactic ambiguities that are usually successfully overcome when further context is introduced. For example, upon encountering the input sentence fragment: *I have…* the agent is not sure whether to comprehend the verb as an auxiliary verb or as the semantic-content verb indicating possession. The next few words usually serve to disambiguate these alternatives in comprehension. On the other hand, since generation is driven by the underlying semantic model, there is no question as to the ambiguity of this construction at the point where the second word is generated. In other words, the generation component is less susceptible to some processing effects that characterize comprehension.

Since the two processes draw upon the same data structures and basic operator architecture, it is also possible to interleave them. Thus, even though massively parallel processing is not available in NL-Soar, a very fine level of swapping between operators of different task types produces the overall impression of the simultaneous global performance of different tasks. Thus several tasks, including language (e.g., comprehension and generation), can be managed 'simultaneously', (e.g., driving a vehicle while speaking on a cell phone) and can be interrupted and swapped (Nelson et al., 1994).

5. ONGOING WORK. In this section we sketch ongoing efforts by the author to extend the basic architecture and capabilities of the system. Current work on generation includes system-specific performance aspects. For example, the generation goal stack is quite deep since comprehension is embedded within generation; this incurs considerable chunking (learning) overhead. Occasional chunks (bundles of learned behaviors) fire out of their canonical order, introducing recognitional state-change sequences that differ from deliberation. A third issue is that, with long and complex sentences, the deep goal stack and attendant numerous state copies sometimes result in memory exhaustion.

NL-Soar comprehension was also upgraded for French processing by adding a French lexicon, French-specific syntactic-model constraints (e.g., permitting object clitics before verbs), and relevant semantic model features (such as gender). This relatively simple process yielded a French comprehension capability. By leveraging the same basic operator-based architecture and the approach outlined above, the development of a French generator was trivial. To date, no other target languages have yet been developed, though to do so would be straightforward.

Specific properties of generation operators are still being addressed. How much semantic material should trigger generation of a piece of sentence structure? How many generation operators should be available at any given time, and how can they be appropriately selected? How specific should the chunks which encapsulate stages of an operator's application be? Should the granularity of generation operators be comparable to those in comprehension?

System improvements are also being made in the area of extending lexical coverage of the system. Currently, the lexicon is largely hand-coded and reflects several hundred entries only. In an effort to scale up the system, an interface has been developed with WordNet (Miller 1993), which provides the system with a much more systematic description of semantic and syntactic properties.

6. CONCLUSION AND ACKNOWLEDGEMENTS. The system described in this paper reflects a general, operator-based approach wherein generation directly implements the comprehension infrastructure; the result is a basic paradigm for integrated, fully compatible language processing using lexical access, syntactic and semantic comprehension, and generation. Interactions across these modalities such as bootstrapped learning, task interleaving, and shared knowledge structures reflect performance capabilities possible via this integration. The system's core approach assures an interesting tension between increased processing potential and cognitively constrained performance. This operator-based generator, which also draws upon analysis operators, is thus a comprehensive language modeling system that exhibits shared functionality and architecture across language modalities.

---

[1] Thanks are due to fellow NL-Soar researchers at CMU (particularly Jill Fain Lehman, Rick Lewis, Julie VanDyke, Nancy Green, Robert Rubinoff, Greg Nelson) and at BYU (Nate Blaylock and Anton Rytting), and to members of the larger Soar project. Any errors are the responsibility of this author.

[2] NL-Soar also integrates a dialogue module that manages discourse turns and their related expectations. This aspect of the system has been documented elsewhere (Green & Lehman, 1996) and will not be discussed here.

REFERENCES

ANDERSON, JOHN R. 1993. *Rules of the mind*. Hillsdale NJ: Erlbaum.
CHOMSKY, NOAM. 1986. *Barriers*. Cambridge MA: MIT Press.
DORR, BONNIE JEAN. 1993. *Machine translation: A view from the lexicon*. Cambridge MA: MIT Press.
GIBSON, EDWARD A. F. 1991. *A Computational theory of human linguistic processing: memory limitations and processing breakdown.* Carnegie Mellon: CMU-CMT-91-125.

GILE, DANIEL. 1990. La traduction et l'interprétation comme révélateurs des mécanismes de production et de compréhension du discours. *Meta* 35.1:20–27.

GREEN, NANCY & JILL FAIN LEHMAN. 1996. *Compiling knowledge for dialogue generation and interpretation*. Carnegie Mellon: CMU-CS-96-175.

JACKENDOFF, RAY. 1990. *Semantic structures*. Cambridge MA: MIT Press.

JUST, MARCEL A. & PATRICIA A. CARPENTER. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological review* 99:122–49.

LAIRD, JOHN E., CLARE B. CONGDON, ERIK ALTMANN & ROBERT DOORENBOS. 1993. *Soar user's manual*, Version 6, Edition 1. EE & CS Department, Univ. of Michigan & SCS, Carnegie Mellon.

LEHMAN, JILL F., RICHARD L. LEWIS & ALLEN NEWELL. 1991. *Natural language comprehension in Soar: Spring 1991*. CMU-CS-91-117.

LEHMAN, JILL F., J. VANDYKE & N. GREEN. 1995. Reactive natural language processing: comprehension and generation in the air combat domain. In *Proceedings of the 1995 AAAI fall symposium on embodied language and action*.

LEWIS, RICHARD L. 1993. An architecturally-based theory of human sentence comprehension. CMU-CS-93-226.

———. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of psycholinguistic research* 25:93–115.

LONSDALE, DERYLE, TERUKO MITAMURA & ERIC NYBERG. 1995. Acquisition of large lexicons for practical Knowledge-Based MT. *Machine translation* 9.

MILLER, GEORGE A. 1993. 5 papers on WordNet. Retrieved December 5, 1999 from the World Wide Web: ftp://ftp.cogsci.princeton.edu/pub/wordnet/5papers.ps.

NELSON, GREG, JILL LEHMAN, & BONNIE E. JOHN. 1994. Experiences in interruptible language processing. *Proceedings of the 1994 AAAI spring symposium of active NLP*.

NEWELL, ALLEN. 1988. *Annotated models*. Unpublished manuscript.

———. 1990. *Unified theories of cognition*.

PRITCHETT, BRADLEY L. 1992. *Grammatical competence and parsing performance*. Chicago: Univ. of Chicago Press.

RUBINOFF, ROBERT & JILL F. LEHMAN. 1994a. *Natural language processing in an IFOR pilot*. Carnegie Mellon: Collected Papers of the Soar/IFOR Project: Spring 1994.

——— & ———. 1994b. Real-time natural language generation in NL-Soar. *Proceedings of the Seventh International Workshop on Natural Language Generation*.

VANDYKE, JULIE & JILL F. LEHMAN. 1997. A process model of learning definiteness in a foreign language. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*. Stanford Univ: CSLI

YNGVE, V. H. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104(5):444–66.

YOUNG, R. M. & R. L. LEWIS. (In press). The Soar cognitive architecture and human working memory. In *Models of working memory: mechanisms of active maintenance and executive control*, ed. by A. Miyake & P. Shah. New York: Cambridge University Press.

GIVEN VS. NEW INFORMATION INFLUENCING CONSTITUENT ORDERING IN THE VERB-PARTICLE CONSTRUCTION

Julia Peters
University of Alberta

INTRODUCTION. Spontaneous oral language production is always remarkable in its unpredictability. A simple picture given to ten different people to describe will undoubtedly generate an equal number of unique utterances comprising different syntactic structures, lexical items, and intonational contours. Within all this variability there must exist certain factors which influence speakers to structure their language as they do; otherwise, we would be in a constant state of indecision about how to encode our message. One way to delve into this broad topic is to look at one particular variable syntactic construction (verb-particle combination) and one potential sequence-influencing factor (discourse information flow).

1. CONSTRUCTION UNDER INVESTIGATION. The variable syntactic structure I chose for this experiment is the English verb-particle construction indicated in the following examples by italics.

(1) *pull up* the anchor, *bang in* the door, *hold off* the enemy

Despite being morphologically distinct, the verb and the particle do form a singular concept, evidenced by the fact that many of the verb particle combinations have one-word near-synonyms, as in (2).

(2) *raise* the anchor, *demolish* the door, *repel* the enemy

The verb-particle construction may often be confused with the V + PP sequence since they both can exhibit the same linear sequence of constituents. However, the verb-particle combination is structurally distinct in that the particle is not fixed in its position relative to the noun phrase (NP). Whereas in (1) the particle precedes the NP object, in (3) the particle follows the NP object with no great change in meaning or effect.

(3) *pull* the anchor *up*, *bang* the door *in*, *hold* the enemy *off*

Since both orderings of constituents can be used to describe the same scene or can generate the same interpretation, it seems obvious that there is more at work than purely semantics in the eventual determination of one order over the other. Of the numerous potential factors which may influence the ultimate sequence, this experiment investigates the role of a particular discourse factor in the course of oral languages production.

2. THE EXPERIMENT.

2.1 MOTIVATION. In 1967, Halliday brought to our attention a tendency for 'given' information (information already known in the discourse context) to precede 'new' information (information not previously mentioned or alluded to) in the linear presentation of a sentence.

If one applies this logic to the sequencing of constituents in the verb-particle construction, it may be that the principle of given vs. new information will influence the distribution of the particle in relation to the object NP. More specifically, if the entity encoded by a particular object NP has not been previously mentioned in the discourse (and would therefore be considered 'new'), would it surface *after* the particle more often than in cases where the entity has already been mentioned, in which case, (being 'given') it would precede the particle?

2.2 METHODOLOGY.

2.2.1. PARTICIPANTS. 56 native speakers of English participated in this experiment. Most were undergraduate students from the University of Alberta, between the ages of 20 to 30, and who had grown up in western Canada. A handful of participants, however, were older or from a different English speaking region of the world. Many individual responses from each participant did not follow the target verb-particle form so could not be used in the analysis.

2.2.2. STIMULI. Participants were presented with simple black and white line drawings of basic agentive events. These pictures were designed to elicit particular target structures in participant responses. The target responses included eight different particles (i.e., *away*, *back*, *down*, *in*, *off*, *on*, *out*, *up*) and nine different verbs (i.e., *pull*, *put*, *hold*, *cut*, *fill*, *shut*, *turn*, *cross*, *blow*). Since the meaning of the verb 'pull' does not explicitly denote directionality, it was used in combination with all of the stimulus particles. See Appendix A for a complete list of target verb-particle combinations, along with the expected NP objects. This set of particles was chosen because they are all productive, frequent, and convey spatial relations that can be depicted relatively easily. Similarly, these particular verbs were chosen because they are fairly frequent and easy to portray graphically. These qualities were important to consider so as to prevent delays that could potentially occur if participants were trying to retrieve an infrequent or unclearly represented word from their lexicon.

Participants were prompted to describe each scene by the textual presentation of the subject NP and the verb at the bottom of the screen. In all target trials, the NP subject written on the screen coincided with the agent of the action so they were all active sentences. All of the agents in the pictures were non-specific people (e.g., *boy*, *woman*, *fish-*

erman, etc.). All of the patients or themes (thus the objects) were common non-human entities whose linguistic labels are morphologically simple and fairly frequent.

In order to prevent participants from thinking about the types of syntactic structures they were producing during the experiment (especially to avoid conscious processing of verb, particle, and noun phrase ordering), I included numerous filler stimulus trials in each section of the experiment. The filler stimulus trials included four scenarios which facilitated dative responses, four scenarios which facilitated passive voice responses, and eight scenarios which facilitated basic transitive sentence responses. (For a complete list of the anticipated filler responses, see Appendix B). Thus, the total number of stimulus item sets was 16 target sets and 16 filler sets.

2.2.3. PROCEDURE. The basic task involved was a picture-driven sentence-completion task. Participants sat in front of a desktop computer and controlled the presentation of stimuli with the space bar of the keyboard. In the first half of the experiment, at the start of each trial, the participants would see the simple black and white line drawing of the event. Two seconds after presentation of the picture, the start of a sentence (NP subject and verb) would appear at the bottom of the screen. Participants were to read aloud the start of the sentence and keep talking to provide a completion that described the picture. Responses were tape recorded and later transcribed. Once they had finished the utterance, participants pressed the space bar to start the next trial.

The second half of the experiment was slightly different in that some additional information was provided to the participants before each trial. Prior to presentation of each picture, two clauses would appear on the screen. Again, the participants would read those clauses out loud, after which they would press the space bar to prompt display of the picture. Two seconds later, the start of a sentence would appear at the bottom of the screen and participants would complete each sentence in the same manner as section one.

The two clauses uttered before each picture-presentation acted to provide the participant with some background context for each picture. In half of the cases, the background context explicitly mentioned the entity that would, hopefully, function as the NP object in the utterance. This prior mention of the object would render it 'given'. In the remaining half of the trials in this section, the background context would not explicitly mention the NP object thus leaving the entity as new information to be introduced in the picture.

To recap, there were three different conditions in the experiment: a no-context condition where no prior linguistic context was given, a new-context condition where the target object was not mentioned in the prior linguistic context, and a given-context condition where the target object was mentioned in the prior linguistic context.

In order to generate data on each picture displayed in each of these three contexts, three separate groups of participants were required so that no one person would see the same picture more than once. The two background contexts were controlled to be, syntactically, as similar as possible for each picture, to prevent some other factor in the background contexts from coming into play. For example, the new-context and given-context clauses for a picture which portrayed a woman blowing up a balloon are given below.

(4) a. New-Context
Clause One: *The woman decided she wanted decorations for the party*
Clause Two: *Before all of the guest arrived…*
Sentence Start: *the woman blew…*
b. Given-Context
Clause One: *The woman decided she wanted balloons for the party*
Clause Two: *Before all of the guest arrived…*
Sentence Start: *the woman blew…*

Notice how clause number one is very similar in both contexts and how clause number two is exactly the same in both contexts. The reason for including the intermediary clause was to introduce a new participant to the discourse so that the speaker would less likely use a pronominal object in the target utterance. Since pronominal objects are, for the most part, restricted to surfacing before the particle, responses including a pronoun would not be revealing or included.

2.3. PREDICTIONS. The no-context condition was included to generate some kind of base-line measure from which to view the results of the contextualized conditions. If no other factors exist which may influence the ordering of the particle and object NP then we may expect a random patterning in the no-context condition, where about half of the responses have the particle preceding the object and the other half following the object. However, because there are many other considerations being processed by the speaker, there may be one ordering which emerges as the preferred one in isolated utterances.

The orderings of the contextualized conditions are easier to predict because the target stimuli are being treated in this experiment so as to influence one particular outcome. If the given vs. new principle is at work in the surface realizations of the particle and object, then one would predict that the object will surface in front of the particle when it is considered given information more often than in cases where it is considered to be new information.

2.4. RESULTS. When considering all the results together, an overall tendency emerged for the verb and particle to be adjacent (with the NP object following). 61% of all utterances (n = 605) followed this order. So, it appears that this principle of ‘verb-particle adjacency’ is an underlying force that other factors need to overcome in order to have any sequencing effect.

The predictions above indicate that verb-particle adjacency would be less frequent in contexts where the NP object is given than in contexts where the NP object is new. This predication is exactly what was realized in this experiment. In the given-context, verb-particle adjacency applied in 54% of the utterances (n = 185), whereas in the

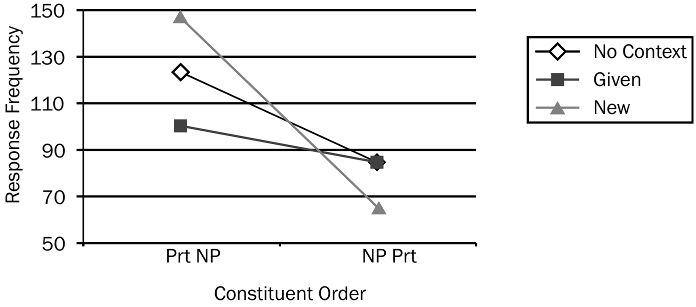


Figure 1. *Response orders across all contexts.*

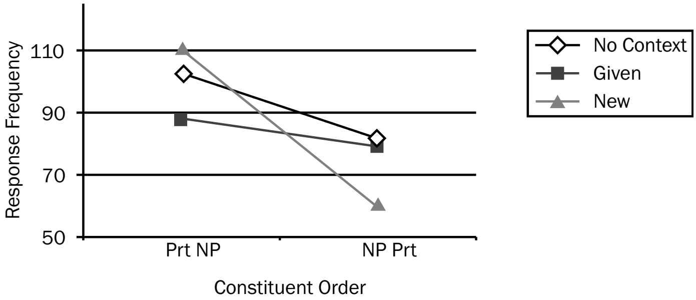


Figure 2. *NP = 2 words responses over all contexts.*

new-context, verb-particle adjacency applied in 69% of the utterances (n = 212). The no-context condition fell in the middle with verb-particle adjacency adhering in 59% of the utterances (n = 208). Figure 1 demonstrates these results in the form of a graph. Notice how the bar representing the new contexts and the bar representing the given contexts intersect at fairly large angles indicating a strong interaction. A chi-square analysis reveals that this interaction is indeed significant ($\chi^2 = 9.819$, $p = 0.002$).

An independent syntactic factor which influences ordering is the length of the NP object. As discussed in Fraser (1975) and demonstrated by Hunter (1981), the more words in the NP object, the more likely this object will be realized at the end of the verb phrase after the particle. Since I did not want this variable to artificially inflate the frequency of verb-particle adjacency, (especially in the predicted new-context), I also analyzed the set of responses that only included NP objects which were two words in length (Det + N). Again, the results agreed with the predictions with verb-adjacency adhering most in the new-context (65%, n = 171), least in the given-context (53%, n = 169), with the no-context responses falling in the middle (56%, n = 185). Figure 2 demonstrates these results graphically with the same basic interactions as are found in Figure 1. As before, the interaction between the new-context and the given-context results are significant ($\chi^2 = 5.270$, $p = 0.02$).

Since the word *pull* had been paired up with all of the different particles, an analysis of the *pull* results may reveal if the specific verb has any affect on the verb-particle adjacency and given vs. new distribution. The results show that the *pull* cases behaved in very much the same way, with verb-particle adjacency applying in 50% of the given-context cases (n = 90) and 65% of the new-context cases (n = 97), another significant result ($\chi^2 = 4.276$, $p = 0.036$).

Although the target particles were defined as that set of eight, in an oral language elicitation task, the range of responses is always greater than anticipated. For this reason, the results included five additional particles (i.e., *open*, *apart*, *aside*, *along*, *aboard*). Along with the anticipated particle *away*, the total number of disyllabic types of particles used was six, with 58 tokens of those types being articulated. An analysis of the disyllabic particles in comparison with the monosyllabic particles revealed an interesting trend. Whereas 65% of the monosyllabic particles (n = 547) were realized adjacent to the verb, only 27% of the disyllabic particles (n = 58) were realized adjacent to the verb. Almost three quarters of the disyllabic particles were uttered after the NP object. A chi-square analysis demonstrates the certainty of this outcome ($\chi^2 = 33.641$, $p < 0.005$).

One last area of analysis to consider is the behaviour of individual verb-particle combinations. For the most part, individual pairings did not have enough individual tokens to produce a statistically revealing analysis. However, some pairings did seem to show certain patterns of behaviour. The combination *cut down* could have been the poster boy for the given vs. new effect camp with verb-particle adjacency only holding in 27% of all the given-context utterances (n = 15) but in 82% of all the new-context utterances (n = 17). On the other end of the spectrum would be the combination *blow up* which seemed to defy influence of given vs. new information with a high 81% of given-context utterances (n = 16) showing verb-particle adjacency. Although this number seems very high, it is still quite a bit lower than the new-context cases where 100% of utterances demonstrated verb-particle adjacency.

One last combination worthy of mention is the *hold back* pairing. Only two out of seventeen people uttered the NP object after the particle as would be the predicted order according to the verb-particle adjacency pattern. Whereas only two speakers said 'hold back the dog', fifteen uttered something like 'hold the dog back'. One possible explanation for this seemingly contradictory result is that *back* is not really a particle in this case but is functioning more like a place adverb indicating the location where the event is taking place (i.e., *back* away from something else). If *back* was an adverb in these cases it would be an adjunct and would, as such, be realized after the NP object complement which is adjacent to the verb. For this explanation to hold water, it must also account for the other *back* data collected for the *pull back* pairing. Although the number of legitimate *pull back* responses is quite low, the results seem to contradict the hypothesis that *back* is an adverb and not a particle. Over half of the total given-context responses (n = 7) have *back* surfacing in between the verb and NP object and almost all of the total new-context responses (n = 6) have the NP object at the end. Therefore *pull back* and *hold back* seem to be behaving quite differently. Arguing that *back* is an adverb in one case and a particle in the other does not seem, to me, to be the best approach. Perhaps, instead, *hold back* is an idiosyncratic pattern where the sequencing of an intervening NP is becoming fixed and resistant to any outside ordering pressures. *Blow up* could also be a case for certain sequences becoming fixed, although this pairing appears to resist separation by an NP object.

3. CONCLUSION. Verb-particle constructions are interesting to the researcher of language production because of the variable ordering of the particle in relation to the NP object. This particular study focused exclusively on oral language production and the interaction of discourse context and the placement of the particle. The most pervasive result was the emergence of the verb-particle adjacency pattern whereby speakers showed a marked preference for the production of the verb and particle next to each other, with the NP object surfacing at the end. Because the verb and particle form a conceptual unit, it is understandable that this conceptual unity would translate into structural unity via syntactic adjacency. However, obviously this structural unity does not hold all of the time or else the verb-particle construction would not be considered 'variable' syntax. Thus, there must exist certain factors which influence alternate orderings or strengthen the frequency of occurrence of the unmarked order. One such factor is the given vs. new property which predicts that if the NP is already given in the discourse context, it will be realized more often in front of the particle than if that NP has not already been mentioned in the conversation. This influence of information flow in a sentence was shown to be at work in many different sets of utterances recorded during the experiment. One set included all of the responses whereas another was controlled for NP object length. The pattern emerged in both of those sets just as when only the *pull* verb combinations were considered. Certain individual verb-particle combinations seemed to display idiosyncratic behaviour whereby they were either very subject to influence by the verb-particle adjacency preference or were quite resistant to this influence. The given vs. new principle obviously plays a small role in the determination of constituent ordering in the verb-particle combination. The multitude of additional factors at work must also be tested and assessed for a better understanding of syntactic production.

APPENDIX A: STIMULI VERBS, PARTICLES, AND ANTICIPATED NOUN PHRASE OBJECTS

PARTICLE	VERB 1	(freq.)	D.O.	VERB 2	(freq.)	D.O.
AWAY	pull	145	the wagon	put	513	the toys
BACK	pull	145	the curtain	hold	509	the dog
DOWN	pull	145	the screen	cut	245	the tree
IN	pull	145	the net	fill	184	the hole
OFF	pull	145	the sticker	shut	50	the light
ON	pull	145	the shirt	turn	566	the radio
OUT	pull	145	the tooth	cross	84	the number
UP	pull	145	the anchor	blow	52	the balloon

APPENDIX B: FILLER SENTENCES (COMPLETIONS PREDICTED)

Transitive

1. The boy kicked the ball over the pylon
2. The boy threw the ball into the net
3. The woman sat in front of the TV
4. The boy ate the carrot at 3:00 PM
5. The woman drove the car into the car wash
6. The girl flew a kite under the power lines
7. The woman closed the door to Dr. Noam's office
8. The boy built a sand castle under the umbrella

Dative

9. The woman read the book to the children
10. The girl mailed a letter to Santa Claus
11. The man offered some cookies to the woman
12. The girl gave a present to her friend

Passive

13. The flower was being watered by the woman
14. The walkman was being listened to by the boy
15. The paintbrush was being used by the woman
16. The letter was delivered by the postman

REFERENCES

Fraser, B. 1976. *The verb-particle combination in English*. The Hague: Mouton.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English 2. *Journal of linguistics* 3:177–274.

Hunter P. J. 1981. *Verb-particle position in English: An experimental study*. MA Thesis, University of Alberta.

◆

DOES SYNCHRONY RECAPITULATE DIACHRONY?
A CASE STUDY FROM JAPANESE

Gary D. Prideaux
University of Alberta

TO WHAT CONSTRAINTS[1] should a synchronic linguistic description (an *analytic* description in Lamb's (1999) terms) be subject? Should it be one which, following Chomsky (1965), captures the native speaker intuitions? Should it be the most parsimonious (ignoring for the moment the question of how this is measured)? Should it be one which captures linguistically significant generalizations? Within generative methodology, it is assumed that these three conditions are all congruent, and a particular description is the 'correct' one if it satisfies all three.

In the earliest proposals offered up by the generative grammarians, it was claimed, albeit implicitly, that the underlying forms of contemporaneous lexical items might lie in their historical antecedents. Could, for example, *mice* be analyzed synchronically in terms of its historical development as a sequence of the Old English form in which /mus/ and its plural /–i/, coupled with phonological rules to account for (a) vowel harmony (/müsi/), (b) open syllable lengthening (/mü:si/ (c) loss of the final vowel /mü:s/, (d) the front vowel unrounding, yielding /mi:s/, and finally (e) the Great Vowel Shift altered /i:/ to /ai/, yielding the modern form /mais/ 'mice'. In such cases, it would appear that the history of the language can 'explain' irregular modern forms. While the *Sound Pattern of English* (Chomsky & Halle 1968) was an advocate of this claim, it was seriously challenged within many studies. Are, for example, *milk* and *lactose* and *galaxy* all related synchronically as well as diachronically? The issue seems in retrospect to be at best amusing and at worst ludicrous.

However, it is precipitous to dismiss out of hand the possibility that historical changes might be directly reflected synchronically within complex morphosyntactic structures. As Hale (1973) pointed out long ago, cases are attested in which sound changes can force the restructuring of entire paradigms. These two possibilities, historical recapitulation and restructuring, taken as extreme poles, allow in principle for competing linguistic descriptions of a particular phenomenon, one of which recapitulates the history of the language while the other reflects a restructuring. Which analysis would be theoretically preferred? Presumably the one which better reflects the native speaker's intuitions, even if that description is less parsimonious.

It is the purpose of this paper to address this problem through an examination of the Japanese adjectival paradigm. Three alternative descriptions of the paradigm are presented, one which is quite transparent, cleaves to the surface and reflects a *historical restructuring* of the paradigm, a second which is quite abstract but which *recapitulates the history of the language*, and a third taken from traditional Japanese school

Gloss	Present	Past	Provisional	Conditional
eat	taberu	tabeta	tabereba	tabetara
wait	matu	matta	mateba	mattara
write	kaku	kaita	kakeba	kaitara
be (inanimate)	aru	atta	areba	attara

***Table 1**. Some verb inflections.*

grammar. Each analysis makes its own predictions as to how native speakers of Japanese process adjectives. A simple experiment was conducted to address this issue, and the three analyses are evaluated against the results of those native speaker judgments.

In what follows, we will first review the Japanese paradigms, and then offer the three analyses. We then turn to the experiment and show that a highly elegant analysis, one which purports to capture linguistically significant generalizations about the language, actually fails to characterize native speaker intuition.

1. JAPANESE VERBS AND ADJECTIVES. In Japanese, adjectives, like verbs, are inflected for tense as well as for several other categories. Japanese is an agglutinating language, rich in sequences of suffixes but relatively impoverished in prefixes. In order to develop the alternative analyses of the adjective paradigm, it is helpful to have before us the verbal paradigm as well. Table 1 provides a selection of verb forms, inflected for a few of the many possible categories.[2]

Japanese verbs are typically partitioned into two classes, those with vowel-final stems (/tabe-/ 'eat') and those with consonant-final stems (e.g., /mat-/ 'wait', /kak-/ 'write', /ar-/ 'be'). The present tense suffix has two allomorphs, /-ru/ after a vowel and /-u/ after a consonant. Various fairly obvious phonological rules are also involved, including two very general and important ones:

(1) k, g → Ø /] i (lenition of velars at a morpheme boundary (]) when /i/ follows)

(2) r → t / __] t (assimilation of /r/ to /t/ at a morpheme boundary)

In addition, it is clear that the two allomorphs of the present tense morpheme (/-ru/ and /–u/) can be reduced to one, /-ru/, by the following consonant-cluster reduction rule:

(3) r → Ø / C] __ (reduction of non-geminate consonant cluster at a morpheme boundary)

This rule expresses the general phonological constraint that the only types of permitted consonant clusters are geminates and nasal + consonant. Thus, the verb paradigm can be expressed in part[3] as in Table 2.

Verbal Inflectional Suffixes			
Present	**Past**	**Provisional**	**Conditional**
-ru	-ta	-reba	-tara

***Table** 2. Some verb inflectional suffixes.*

Gloss	Present	Past	Adverbial	Provisional	Conditional
tall	takai	takakatta	takaku	takakereba	takakattara
cold	samui	samukatta	samuku	samukereba	samukattara
large	ookii	ookikatta	ookiku	ookikereba	ookikattara

***Table** 3. Some adjective inflections*

It is, however, the adjectival paradigm which is our present concern. An example of the ways in which the adjective can be inflected is found in Table 3.

The questions at the heart of this paper are: What is the best way to treat the adjectival paradigm? Should adjectives and verbs be treated separately, each with its own distinct set of inflections, or should they be treated within a single system of inflections? The empirical issue is: How do native speakers of Japanese conceive of this paradigm? It is to a comparison of different analyses of the adjective paradigm that we now turn.

2. COMPETING ANALYSES. Many descriptive linguists who have worked on Japanese, such as Jordan (1963), treat verbs and adjectives as two distinct lexical categories, each with its own inflectional paradigm, while others have tried to collapse the two into a single category. The following three analyses are broadly representative of the families of analyses which have been proposed. The first maintains the distinction between verb and adjective inflections, while the second does not. The third is a traditional Japanese grammatical analysis.

2.1. CONCRETE ANALYSIS. The most common description offered for the two paradigms is one which is both transparent and relatively simple. It treats verbs and adjectives as separate lexical classes with the stipulation that both are marked for virtually the same set of categories, though with different allomorphs for each. Within this analysis, adjective stems are all vowel-final (e.g., /taka-/ 'tall'), and the present tense marker is /–i/, while the past tense marker is /–katta/. According to this analysis, the adjective should be parsed as in (4).

(4) ADJstem + suffix (e.g., taka+katta > /takakatta/)

This highly transparent analysis is summarized in Table 4. An important feature of this analysis is that each suffix is monomorphemic, each with its own unique semantic content.

Adjectival Inflectional Suffixes				
Present	**Past**	**Adverbial**	**Provisional**	**Conditional**
-1	-katta	-ku	-kereba	-kattara

Table 4. *Concrete adjective analysis.*

Adjectival Inflectional Suffixes				
Present	**Past**	**Adverbial**	**Provisional**	**Conditional**
-Ø	-ar-ta	-u	-ar-eba (=ereba)	-ar-tara

Table 5. *Abstract adjective analysis.*

This analysis makes the empirical prediction that native speakers will treat the stem as vowel-final and constant throughout the paradigm, and that each of the suffixes is monomorphemic.

2.2. ABSTRACT ANALYSIS. Since verbs and adjectives participate in almost identical paradigms, since they function identically in main and subordinate clauses, and since the only other inflected category in the language is the copula, it is tempting to formulate a single unifying analysis encompassing both verbs and adjectives. Such an analysis has been proposed by, among others, Makino (1968). This typical generative analysis is both highly parsimonious, since it treats verbs and adjectives in a unitary manner, and also to a large extent it recapitulates the historical development of the adjective paradigm (Chamberlain 1907, Henderson 1945, Martin 1987, Miller 1968).

The essence of this analysis is that adjective inflections are derived from a sequence of an adjective stem and an inflected form of the existential verb /aru/. Such an analysis clearly mirrors the historical development of the paradigm, as Henderson (1945), Martin (1987), and others have shown. However, this analysis requires that the adjective have a stem-final /k/, which is common throughout the paradigm except in the present tense. As Chamberlain (1907) and Miller (1968) both observe, the /k/ is an adjective class marker in Old Japanese. According to this analysis, then, any adjective should, in all cases but the present tense, be parsed as in (5).

(5) ADJstem + ar + suffix (e.g., takak+ar+ta > /takakatta/)

According to this more abstract but historically transparent analysis, the inflectional suffixes for the adjectives in table 3 should be treated morphologically as in Table 5.

But how is the most frequent and common form, the present tense of the adjective, to be handled? Makino's (1968) solution is that the present tense, the most unmarked member of the paradigm, has no suffix at all. Thus, the present tense form is simply the stem, e.g., /takak/). But a completely general phonological constraint of Japanese is that a word may not end in a non-nasal consonant. The phonological resolution of this problem is the addition of an epenthetic vowel, typically /-i/, after the final (non-nasal) consonant, as represented in rule (6).

(6) Ø → i / C__# (the epenthetic vowel /i/ is added to a word-final consonant, where # is a word boundary)

Thus, the present tense (unmarked) form for an adjective would simply be:

(7) ADJstem # (e.g., /takak#/)

And to this stem would be added the epenthetic vowel /i/ (by rule (6), yielding:

(8) ADJstem - i # (e.g., / takaki#/)

We recall the general rule of velar lenition, the deletion of a /k/ or /g/ at a morpheme boundary before a high front vowel (rule (1)). The derivation of a present tense form of an adjective like /takai/ 'tall) would then be :

(9) takak# > takaki# (rule 6) > /takai/ (rule 1)

Thus, this analysis, while quite abstract, nevertheless recapitulates the historical development of the adjective paradigm. Moreover, it is parsimonious to the extent that it unites adjectives and verbs into a single analysis

This analysis, like the first, makes an empirical prediction. It predicts that native speakers of Japanese recognize that an inflected adjective consists of a stem (ending in /-k/) plus an inflected form of the verb /aru/, with virtually all the suffixes being polymorphemic.

2.3. TRADITIONAL ANALYSIS. A third analysis should also be explored here, namely that of the traditional Japanese school grammarians (see Chamberlain 1907, Ikeda 1980). This analysis is highly detailed and treats verbs and adjectives as inflected forms, with basically the same categories for each, in spite of the fact that the suffixes are quite different. However, this traditional analysis treats every member of the adjectival paradigm as having its own unique stem and unique suffix. One reason for this complexity derives from the Japanese writing system which is composed of *kanji*, borrowed and adapted Chinese characters (used primarily for the meaning-bearing roots of nouns, verbs, and adjectives) and syllabaries (used for inflections, function words, and the like). The syllabary used for inflections (*hiragana*), like the other (*katakana*) system, does not permit the writing of a single consonant other than a nasal or a geminate cluster. It is impossible to have within the traditional Japanese writing system a break between all phonemic segments which an alphabetic system allows. Since only syllable (or mora) boundaries are expressible, it would not be possible to write a form such as /takaku/ as *takak-u* but only as *taka-ku*.

The traditional Japanese analysis of the adjective forms is primarily taxonomic in nature and is exemplified in part in Table 6, for the adjective /takai/ (after Henderson

Adjectival Inflectional Suffixes				
Stem	**Suffix**	**Gloss**	**Traditional Japanese label**	
taka	-i	present tense	*rentaikei* *shuushikei*	'prenominal form' and 'terminal form'
takakat	-ta	past tense	*renyookei*	'adverbial form'
takaku	-naru	adverbial	*renyookei*	'adverbial form'
takakaro	-o	tentative	*mizenkei*	'incomplete form'
takakere	-ba	provisional	*izenkei*	'conditional form'

***Table 6**. Traditional Japanese adjective analysis.*

1945). As Table 6 reveals, each member of the paradigm has its own unique stem and its own suffix.

The traditional analysis also makes empirical predictions as to how native speakers will parse their adjectives, namely that there will be different stems for each member of the paradigm, plus distinct suffixes, each uniquely keyed to its stem allomorph. This is highly complex, ad hoc, and presents a quite awkward and counter-intuitive analysis. It does not reflect any overall generalizations nor does it respect obvious candidates for morpheme boundaries.

In order to test these three quite distinct analyses of the Japanese adjectival paradigm, a simple experiment was carried out. In it, native speakers of Japanese were asked to parse Japanese verbs and adjectives into as many components as they considered relevant. It is to a discussion of this experiment that we now turn.

3. THE EXPERIMENT. The 26 volunteer participants in the experiment were 17 females and 9 males, ranging in age from 17 to 59, though the vast majority were in the 30–50 age range. All had received their pre-university schooling in Japan. Both the instructions and the stimulus materials were provided in Japanese. Since, as we saw above, the Japanese writing system uses both *kanji* and two distinct syllabaries, it is not possible to separate all phonological segments within a syllable. Thus, the stimulus materials were presented in both the native orthography and in a romanized version (*romaji*). The stimulus materials contained a set of randomly ordered conjugated verbs and adjectives. In all, 22 different adjectives were used, and while not all of the many forms of any adjective or verb were included, those cited in Tables 4 and 5 above were all included. Each of the five paradigm members represented in Table 2 was represented with eight tokens, yielding a total of 40 adjectives in the stimulus set. Distracters were also included. Participants were instructed about morphemes as meaning units and were given examples (taken from nominal, not verbal or adjectival, constructions). They were then asked simply to divide each stimulus item into as many meaningful units as they thought were present. The task was not onerous and took only about 20 minutes to complete.

Suffix									
Present			Past				Adverbial		
0	i	other	katta	kat-ta	atta	ka-tta	ku	k-u	0
139	67	2	157	26	8	17	172	8	28

Table 7. *Frequency of parses for present, past, and adverbial forms.*

Suffix										
Provisional					Conditional					
kereba	kere-ba	ke-reba	ereba	reba	kattara	katta-ra	kat-tara	attara	kat-t-ara	ka-ttara
117	71	2	8	10	70	57	41	8	8	24

Table 8. *Frequency of parses for provisional and conditional forms.*

4. RESULTS AND DISCUSSION. The patterns found in the data were not random, but rather interesting and informative patterns emerged. The tabulated results for the present, past, and adverbial data are found in Table 7. These forms are highly frequent in the language. The provisional and conditional data, far less frequent both in conversational and written language, are found in Table 8. We examine each set separately.

For the present tense, by far the majority of participants treated a simple adjective such as /takai/ as having no suffix at all, in spite of the fact that the word is written in Japanese as a single *kanji* for the root /taka-/, followed by the syllabic /i/. While there is not complete unanimity amongst the participants, there is a strong statistical tendency for the present tense to be viewed as monomorphemic ($\chi^2 = 135.5$, $p < .00001$) in contrast to both traditional and the concrete analyses, both of which treat /–i/ as a tense marker suffix. It should be noted that for the present tense, however, the 'runner-up' is /-i/. Thus, for the present tense, the abstract analysis wins.

The past tense form also shows considerable variation, but the dominant preference across the participants is for /-katta/, a single morpheme marking the past tense ($\chi^2 = 285.8$, $p < .00001$), with a distant second the form /kat-ta/, which is predicted by none of the analyses, though the abstract analysis does predict a morpheme boundary within this form, so this provides some support for the abstract analysis as the runner-up. Finally, the adverbial preference is for /-ku/ as the suffix ($\chi^2 = 230.9$, $p < .00001$), though a goodly number of participants chose to view this as monomorphemic. Here the concrete analysis wins handily, with the abstract and traditional analyses representing the runners-up.

For the first three, highly frequent, forms of the adjective conjugation, it appears that none of the three analyses is uniformly supported. In particular, the more concrete analysis is supported for the past and adverbial forms, while the more abstract analysis seems supported for the present tense. Or is it? We return to this below, but first let us examine the other two forms of the adjective which were investigated here.

For the less frequent forms, those found in table 8, we find much more scatter and variation, as might be expected. For the provisional member, the clear winner ($\chi^2 = 246.3$, $p < .00001$) is the monomorphemic suffix /-kereba/ (supporting the

Paradigm	**Preferred form**	**Runner(s)-up**	**Predicted by category analysis**
present	Ø		abstract[4]
		-i	concrete (and traditional)
past	-katta		concrete
		-kat-ta	abstract
adverbial	-ku		concrete
		Ø	abstract (and traditional)
provisional	-kereba		concrete
		-kere-ba	abstract (and traditional)
conditional	-kattara		concrete
		-katta-ra	abstract
		-kat-tara	abstract (and traditional)

***Table 9.** Overall results for the adjective paradigm*

concrete analysis), while the runner-up is the bi-morphemic /-kere-ba/, predicted in part by both the abstract and traditional analyses, since both predict an internal morpheme boundary. The conditional form shows the most variability in the distribution of its forms, but the apparent winner is /-kattara/, predicted again by the concrete analysis. There is a statistically significant difference among the three top contenders ($\chi^2 = 24.8, p < .0001$) , although not between the two runners-up, /katta-ra/ and /-kat-tara/ ($\chi^2 = 2.61, p = 0.25$.). However, both are bi-morphemic, which could give some limited support to the abstract analysis.

These overall results are summarized in Table 9. As the table reveals quite clearly, the overall preferred analysis, the one which the participants in this study seemed to follow, is the more concrete analysis, with the abstract and traditional analyses playing some role in the 'runners-up' results. What are we to make of this?

The fact that there is scatter in the data, especially among the relatively less frequent forms, suggests that a profound restructuring of the paradigm is well underway. The restructuring appears to have been pretty much complete in the more frequent forms (Table 7). Both the past and adverbial forms clearly represent a restructuring, in which the old stem-final /k/ has been reassigned to the suffix, and the internal structure of the past tense morphology has been erased, leaving simple monomorphemic suffixes. Restructuring is less evident for the infrequent forms, but the tendency is clearly underway, having advanced further in the provisional than in the conditional forms. A guiding principle seems to be that the entire system is gravitating to monomorphemic representations. That is, the more abstract analysis, which recapitulates the history of the language, is being replaced by a more transparent, concrete analysis for contemporary native speakers.

While it might appear that the present tense form is an exception to this generalization, I do not think that is in fact the case. Being the most frequent form of the paradigm, but with the stem-final /k/ absent in it, though present in all the other forms, the present tense seems the most fossilized in its monomorphemic status, a view reinforced by the traditional grammatical analysis. Thus, it may well be the most advanced member of the paradigm in the restructuring process.

The less frequent forms, the provisionals and conditionals, appear to be the most unsettled members on the road to restructuring. The internal, historically valid, morpheme boundaries are giving way under pressure for monomorphemic status throughout the paradigm, but the tendency, especially within the conditionals, appears to be influenced strongly by issues of analogy. The conditional, after all, does share with the past tense form the sequence /katta/, so it is not surprising that the native speaker, especially in a task like the present one, would see that parallel and exploit it. Nevertheless, the monomorphemic form does win out, and the move toward monomorphemic adjectival suffixes seems well on the way to completion.

5. CONCLUSIONS. Three major conclusions can be drawn from this study. The first is that no matter how elegant and parsimonious a particular linguistic analysis might be, it does not necessarily explain the behavior of native speakers. If it means anything at all, 'linguistic competence' must tell us about how actual native speakers really do use their language and it must reveal empirically what they know and what they do. The abstract analysis, while elegant and appealing for its simplicity, fails in that task.

Secondly, restructuring is taking place, and a guiding force seems to be semantic transparency at the surface. There is no reason for a native speaker to retain old morphemic boundaries, especially when complex sound changes (such as the lenition of velars before high vowels or the gemination of unlike consonants to like ones) take place. These sound changes obscure the earlier morpheme boundaries, as is clearly the case in the examples we have examined above.

Thirdly, a cautionary note is in order. The experiment reported here does employ native speakers in an analytic task. It is possible, as it always is in experimental studies, that the nature of the task itself is driving the results. In order to shore up the conclusions offered here, it is necessary to carry out further and more refined studies. It might, for example, be useful to exploit more sensitive on-line methods such as lexical decision or phoneme monitoring tasks to try to get at the same phenomena. So, at this point, while our results are all statistically significant and quite compelling, they are still preliminary.

Yet the question posed in the title of this paper seems to be answered, at least in the context of the Japanese adjectival paradigm: synchrony did at one point recapitulate diachrony, but restructuring is well underway. And the best evidence is that the most revealing analysis of the Japanese adjective paradigm, reflected in the behavior of native speakers, is that the historical morpheme boundaries within the suffixes are being erased in favor of a more transparent, surface-oriented analysis.

---

[1] In this paper, I return to a topic which I first attacked, rather unsuccessfully, some years ago (Prideaux 1976). I am grateful to Ms. Satomi Currah and Dr. Kaori Kabata for their help, as well as to the volunteer Japanese speakers who took part in the experiment. I

am also grateful to the anonymous referees who suggested important improvements. This research was supported in part by Social Sciences and Humanities Research Council of Canada Research Grant 410-96-0349.

[2] For convenience, I follow the grammatical nomenclature used in Jordan (1963).

[3] Of course, Table 2 does not exhaust the verbal paradigm, since there are numerous other members. It is rather intended to be illustrative of the full paradigm.

[4] As an anonymous reader pointed out, it might be argued that this is really a concrete analysis, not an abstract one, since the speakers might arguably have adopted a completely suppletive allomorphy here. This possibility is elaborated briefly below.

REFERENCES

Chamberlain, Basil Hall. 1907. *A handbook of colloquial Japanese*. London: Crosby, Lockwood & Sons.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

——— & Morris Halle. 1968. *The sound pattern of English*. Cambridge MA: MIT Press.

Hale, Kenneth. 1973. Deep-surface canonical disparities in relation to analysis and change: An Australian example. *Current trends in linguistics* 11:401–58.

Henderson, Harold B. 1945. *Handbook of Japanese grammar*. London: George Allen & Unwin.

Ikeda, Tadashi. 1980. *Classical Japanese grammar illustrated with texts.* Tokyo: Tookoo Gakkai.

Jordan, Eleanor Harz. 1963. *Beginning Japanese*, 2 parts. New Haven: Yale University Press.

Lamb, Sydney M. 1999. *Pathways of the brain: The neurocognitive basis of language.* Amsterdam: John Benjamins.

Martin, Samuel E. 1987. *The Japanese language through time*. New Haven: Yale University Press.

Makino, Seiichi. 1968. *Some aspects of Japanese nominalizations*. Tokyo: Tokai University Press.

Miller, Roy Andrew. 1968. *The Japanese language*. Chicago: University of Chicago Press.

Prideaux, Gary D. 1976. An excluded generalization. *In Linguistic and literary studies in honor of Archibald A. Hill: Volume 1: General and theoretical linguistics*, ed. by Mohammad Ali Jazayery, Edgar C. Polomé & Werner Winter, 245–55. Lisse: Peter de Ridder Press.

~

CATEGORIZING ENGLISH SUPERPOSITION

Bernard Paul Sypniewski
Rowan University, Camden Campus

Superposition is a general theory that does not apply to any particular language but applies to all languages. Although it is part of Applicative Universal Grammar (AUG) Shaumyan (1987), superposition should interest all linguists because it provides a framework into which many other theories, such as nominalization and grammaticalization can be placed. By using superposition rather than notions from several disparate theories to explain complex linguistic behavior, we can simplify our theories without any loss of rigor. When seen from the point of view of superposition theory, typology will be enriched, extended and made more flexible. This paper is not an attempt to convert linguists to AUG. Rather, it is part of a study of a general process that underlies a number of previously identified forms of linguistic behavior. One need not be an adherent of AUG to use superposition theory.

Briefly, in superposition, a secondary function[1] of a linguistic expression[2] *E* is overlaid upon its primary function[3]; Shaumyan (1987), Sypniewski (1997), creating a new, bistratal function that allows *E* to perform some aspects of both primary and secondary functions simultaneously. Superposition occurs when the context requires it to occur. A *superposer* operates on a *superponend*. A *superposer* is that linguistic entity or device that causes a superponend to be superposed. For example, the superposer *-ing* operates upon the superponend *turn* to produce *turning*. The superposer overlays the secondary function of term on the primary function of predicate, e.g., *the turning of the seasons*. As this paper shows, a superposer may be but need not be a portion of a word such as an affix.

Superposition gives language a simple, flexible, and robust tool for extending the lexicon. For example, although *–ing* in the example above superposes a secondary function of term on a primary function of predicate in *turning*, the same superposer may superpose a different function on the same word in a different context, e.g., *the turning wheel*. Superposition operates within a given context and is dependent on that context for the result that it produces. Superposition is not the only mechanism used to expand the lexicon, of course, but superposition is one of the common mechanisms used to create new expressions. Superposition is conservative; it uses already existing linguistic material to form new expressions. Superposition, by itself, does not create permanent entries into the lexicon. Superposition occurs only within a limited context, i.e., within a single expression. Superposition reoccurs every time that the expression is repeated. If a word or expression *W* is not repeated a sufficient number of times by a sufficiently large portion of the linguistic community, *W* will not become part of the permanent lexicon. On the other hand, if an expression meets the needs of the linguistic community, it will be repeated often enough to become part of the permanent lexicon[4]. Thus, it is the combination of superposition and repetition that creates new entries into the lexicon. Even when the expression becomes part of the permanent lexicon, the superposition can still be seen. Superposition is not only implicated in changes in word function but also in changes in word meaning[5] (see Sypniewski 2000). Most of the examples in this paper are examples of functional superposition. However, the semantics of an expression can also superpose an expression (ibid).

Superposition is not just another name for nominalization or grammaticalization. If anything, nominalization and grammaticalization are subsets of the theory of superposition. According to AUG, any expression may be superposed, not just nouns or verbs. While it is beyond the scope of this paper, it is likely that expressions can be superposed to perform functions beyond those of the traditional linguistic categories. This paper describes the simple, generalized mechanisms used in the process of superposition in English. Other languages may not use superposition as freely as English or may not use both types of superposition mentioned in this paper to the same degree or in the same way as English. There has been some work done on nominalization in Thai that may prove useful for the understanding of superposition Prasithrathsint (1994); Prasithrathsint (1997). New abstract nominal expressions have developed in Thai within the last century. Some scholars believe that these new nominal expressions result from Thailand's exposure to the West and the increasing familiarity of Thai speakers with English. What is of interest for us is that the new Thai nominalizations use material superposers, e.g., /kann/ and /khwaan/ are garmmaticalized morphemes or prefixes. They derive from terms meaning 'affair' and 'matter', respectively[6]. Prasithrathsint's interesting studies demonstrate that semantics (the reasons why the nominalizations occurred) of the host language (Thai) expressions may be intended to model the semantics of another language (English) but the superpositions occurred in a manner appropriate to the grammar of the host language.

Superposition may allow us to categorize languages. We can create a continuum with languages that use only material superposition on one end and ones that use only contextual superposition on the other. A language L[7] appears on this continuum depending on some measure of the utility of both types of superposition to the speakers of L. This method of grouping languages may provide very different results than the traditional 'family of languages' since such a continuum would relate languages by an aspect of their functional similarity rather than by their phonetic similarity.

1. MATERIAL AND CONTEXTUAL SUPERPOSITION DEFINED. There are two broad categories of superposition, which I shall call *material superposition* and *contextual superposition*. Superposers from both categories may operate in tandem in the same linguistic context but no individual superposer is a member of both categories. *Material superposition* overlays a secondary function onto the primary function of a word[8] W due to some change in, addition to, or subtraction from, the structure of W. The sign indicating the structural change is the superposer[9]. *Turning* is an example of material super-

position because it involves the addition of a suffix to a root; the suffix operates on *turn* to produce *turning*. In English, suffixes are the most obvious but not the only material superposers. *Structure* refers to any aspect of the morphology, phonology, stress, accent, pronunciation, inflection, intonation[10], or writing of *W*. The differences in the pronunciation of the words *minute* and *arithmetic* in (1) serve to superpose the secondary function of term modifier onto the primary function of term in the second instances.

(1) a. minute by minute/a minute change
    b. arithmetic/arithmetic function

Contextual superposition overlays a secondary function upon the primary function of *W* in a linguistic expression *E* by the change of, addition to, or subtraction from, *E*. *W* need not be altered at all. In the first instance of Example 1, the word *minute* has been placed before *change*[11], as though it was a term modifier. In the phrase, *Zoo Run*[12] the context imposes a secondary function of term modifier on *Zoo* and a secondary function of **term** on *Run*. Neither word is altered in any way. The context[13] is the superposer in this example.

Material superposition is internal to *W*. Contextual superposition is external to *W*. Material superposition affects a single word. Contextual superposition may affect more than one word at a time. In our *Zoo Run* example, one contextual superposer affects both words. We may say that the contextual superposer is the placing of the word *Zoo* before the word *Run*. A *context* is the relation of each word in a linguistic expression to the expression as a whole. Every linguistic expression must have at least one word or, in spoken language, one utterance. The one-word sentence *Stop!* is complete as both a sentence and a linguistic expression. *Stop* is a predicate. AUG postulates that all sentences (s-types) must have something that performs a predicate function. The tone of voice, if uttered, or the exclamation point, if written, reinforce the notion that *stop* is a command, hence a complete sentence. We may say that the context is the fact that *stop* is the sole word in the sentence or that *stop* is accompanied by strong inflexion and, perhaps, loud tone. The context superposes the secondary function of **s**[14] on *stop*'s primary function of predicate.

An expression may use more than one incidence of superposition. In *running shoe*, *running* is doubly superposed. *Running* is materially superposed by the suffix *–ing*, overlaying the secondary function of term on *run*'s primary function of predicate. *Running* is contextually superposed by occurring before or being spoken before the term *shoe*. The context overlays the function of term modifier on the already superposed *running*. We may think of all instances of superposition in a linguistic expression as occurring simultaneously or as occurring one after the other, like the composition of functions in mathematics. In any event, our brains process language so quickly that all instances of superposition in an expression are, for all practical purposes, simultaneous.

2. A BRIEF DESCRIPTION OF AUG'S TYPE SYSTEM. Words have types because of the functions that they perform in a context[15]. Words have primary, secondary, and even tertiary types. The primary type of *W* is the function that *W* most frequently performs. Words are not the only linguistic objects that have types. An entire expression has a single type[16]. For example, what is traditionally thought of as a noun phrase is, for AUG, an expression with the type t or term. AUG describes two broad functional categories on the genotype level[17]: operators and operands. *Operator* is a general word that refers to a linguistic function. An *operand* is a linguistic entity on which an operator operates (called an *operation*). Every operation produces a result, called the *resultant*.

Every expression is an example of the *nucleus::satellite* relation. Every expression has a nucleus and may have one or more satellites. The nucleus of a nucleus::satellite relation is a term for expressions whose type is term and predicate for expressions whose type is either predicate or s-type (a sentence). The function of predicate is to be the nucleus of a sentence. If a word or expression *F* functions as the nucleus of a sentence, *F* has the type[18] predicate regardless of whether *F* has the primary type of predicate. A noun phrase[19] has a term as its nucleus. If a word or expression *K* functions as the nucleus of an expression with the type term, *K* has the type term regardless of *K*'s primary type. A term serves the function of being an operand for a predicate, and, in that sense, is the satellite of a predicate. From this discussion, we see that if *F* does not have the primary type of predicate, *F* is superposed to predicate. If *K* does not have the primary type of term, then *K* is superposed to the type term. Material superposition, therefore, operates by changing the *nucleus::satellite* relation.

(2) *The Zoo Run raised $10,000 for the Philadelphia zoo.*

In (2), the first instance of the word *zoo* is superposed by being placed in a satellite relation with the word *run*. Normally, *zoo* is a term and, therefore, the nucleus of an expression with the type term. *Zoo* is superposed because it is the satellite[20] of *run*. *Run* is superposed because it is the nucleus not of a sentence but of a term. The word *Philadelphia* is contextually superposed because it is the satellite of the second instance of *zoo.* (2) demonstrates that superposition is 'local', i.e., superposition operates within a narrow context. While the first instance of *zoo* in Example 2 is superposed, the second instance of *zoo* is not affected by the superposition. For the purposes of this discussion, we will assume that 'context' means 'nucleus::satellite relation'. Superposition operates on expressions rather than words even if the expression consists of one word. Several instances of superposition may occur at once, e.g., *Zoo Run* or *running shoe*. AUG claims that the types of words and expressions are dynamic. Most linguistic theories assume that the types of words and expressions are static.

Some words seem easier to superpose than others are. A good example is the word *run*. Some examples of the superposition of *run* are found in Table 1.

Functional superposition (Sypniewski 1997) often affects the semantics of a linguistic expression and *vice versa* (ibid 2000), as can be seen in Table 2 and in the

1	The boy *ran* to the store.	No superposition
2	The *running* shoe	Material superposition (-*ing*) and contextual superposition
3	The Battle of Bull *Run*	Contextual superposition
4	A *run* in her stocking	Contextual superposition
5	No hits, no *runs*, no errors	Contextual superposition (and material superposition (-s)?
6	The *runner*	Material superposition (-er) (and contextual superposition?)
7	Who *runs* London?	Semantically superposed but not functionally superposed

Table 1.

1	*running* shoe	*running* joke
2	*run* of the mill	*run* in her stocking
3	the *runner* on the field	the *runner* on the steps
4	*running* the government	*running* the sheep
5	the engine is *running*	my nose is *running*

Table 2.

seventh example in Table 1. While structurally similar, are these pairs of expressions semantically the same?

Semantic superposition complicates the theory of superposition. Semantic superposition is not a completely worked out theory but, at this stage, it is sufficiently well-developed to say that the general theory of superposition cannot be understood without taking semantic effects into consideration. Table 1 (except for #1) contains examples of functional superposition. In Table 2, there is no functional change between members of the same pair but sometimes the semantic difference is substantial, as in #3 and #4. In #3, the first phrase refers to a person and the second phrase refers to an object (a rug of some sort). In #4, the first phrase refers to operation; the second phrase refers to physical movement or managing physical movement.

3. USES OF SUPERPOSITION. We may see the uses to which English speakers put superposition by examining some difficult-to-superpose words. A good example is *antihistamine*. *Antihistamine* is a term; the word refers to a class of drugs. All terms are easy[21] to superpose to term modifiers. *Antihistamine* is no exception: *antihistamine tablet*. It may be difficult to see how *antihistamine* can be superposed any further. Unlike *run*, which can be easily superposed to a term, *antihistamine* does not describe any activity or action[22] that would allow us to easily superpose it to a predicate. However, someone might say: *The doctor antihistamined me*. This is certainly not an ordinary use of *antihistamine* but the sentence is understandable, perhaps requiring a moment to grasp its meaning[23]. The sample sentence can refer to a doctor's providing or suggesting antihistamines to a patient for some ailment. I do not suggest that this use will ever become an accepted part of the general lexicon, but the sample sentence demonstrates that superposition may be used in the creation of new words or phrases. Here are four more examples of similar linguistic behavior:

(3) a. New Jersey Network is proud to partner with the Division of Youth and Family Services.
b. Kinko's: it's a new way to office.
c. We repurposed the data
d. XYZ Corporation took over ABC Corporation in a leveraged buyout.

In Example 3, terms (*partner*, *office*, *(re)purpose*, and *leverage*) are used as predicates or, in the case of *leverage*, contain aspects of the predicate in their new roles. Superposition frequently occurs in the creation of new words or phrases. A language group often creates new words out of compounds of old words. The lexicon greatly expands by the compounding of old words into new words or expressions that are used as one word. Common animal and plant names form a large group of words made in this way. Common or 'folk' names are generally deprecated by the scientific community, yet they show a remarkable tendency toward describing the community's idea of the important characteristics of a plant or animal. The modifier *horse* in (4)e and (5)e refers to the large size of the chestnut or reed or, more specifically, to their flowers or foliage. The *skunk* in *skunk cabbage*, (4)d, refers to the delightful aroma of the plant.

(4) a. Lightning bug
b. Gypsy moth
c. Honey Bee
d. Skunk cabbage
e. Horse chestnut

(5) a'. Stink bug
b'. Luna moth
c'. Carpenter Bee
d'. Skunk
e'. Horse tail

(4) can be extended almost indefinitely. Superficially, the members of (4) can be thought of as two-word phrases, but each entry is one name. Substitute or remove a word and a different animal or plant comes to mind, as in (4). All of the examples in (3) and (4), except for (4)d, are examples of contextual superposition. (3)a and b also contain examples of material superposition. Many common words are, in reality, two words used as one, e.g. *afternoon, alleyway, basketball, hatbox*, and *pitchfork*.

I extracted[24] a subset of 250 compound words from a corpus of about 1500 examples of superposition. In building my corpus of superposition examples, I have tried to get a random selection of authors and types of texts so as to minimize the impact that the peculiarities of any authors or text types might have on the conclusions that I might draw. By considering each compound word as two words that happened to be joined together, I examined the type of the first word, the type of the second word, and the resultant type. If this small sample is typical of compound words, superposition is heavily skewed towards the use of and production of terms. This finding is not unexpected since we use language to describe and symbolically manipulate the world. Terms allow us to refer to parts of our experience in a sort of shorthand. Terms are easy to superpose into term modifiers, e.g., *ship* (term), *ship mate* (term modifier), and frequently into predicates, e.g., *to ship* (predicate)[25]. Superposition allows us to group concepts. We can use superposers to describe an object *O*, some aspect of the existence of *O*, and some action performed with or to *O* without substantially expanding the lexicon. In this sense, superposition is a conservative mechanism that, on one hand, increases the scope of the lexicon by increasing the flexibility of expressions, and, on the other hand, keeps the size of the lexicon manageable.

In some types of writing, superposition occurs quite frequently. One text that I selected was Gilbert Waldbauer's *Insects Throughout the Seasons*. I isolated a number of examples of compound nouns. In one fifty page segment of Waldbauer's book, I counted 213 compound words that resulted from superposition. At an average of 370 words per page, superposed compound words amounted to 1.32% of the text (213/18500). One may argue that compound words like *seahorse* or *highrise* are individual words and not, therefore, superposed. One must, however, acknowledge that there is no difference between *sea horse* and *seahorse* or *high rise* and *highrise*. At one time, these words were superposed. Here, the superposer is the context that associates two words as parts of a single name. In each case, as in (4) and (5), a change in either part of the compound changes the semantics of the entire expression. That many compound words are spelled as one word rather than as two is merely the recognition of their use as a single name. These examples, and from experience we can think of hundreds more like them, demonstrate the role of superposition in expanding the lexicon. If a superposed expression E results in E_s and E_s proves useful to the linguistic community, E_s may be adopted as a word or expression in its own right. Will *office* become a predicate as in (6)? I have not heard it used as a predicate except for the one commercial mentioned. What about *partner* or *repurpose*? Certainly, they are used more often. Many words go in and out of fashion. Fashion or necessity is determined by the linguistic community. Superposition is the mechanism that allows the free creation of new, temporary[26], expressions. It is often easier to 'extend' a meaning than to create a new expression from whole cloth. We often hear the results of superposition in

(6) a. slang, e.g., *offed* (slang for killed), *witch hunt, a done deal*
  b. technological jargon, e.g., *to upgrade software, disk cache, object oriented programming*
  c. fashion, e.g., *stiletto heels, tailored twills, flirty cocktail dress*
  d. entertainment, e.g., *blonde bombshell, blockbuster*[27]*, hiphop*
  e. sports, e.g., *a three-pointer, round-robin, dugout, touchdown, round-tripper, shot put, fly out*
  f. publishing, e.g., *book jacket blurbs, Martha Stewart Living*

The sentences in (7) were culled from recent publications:

(7) a. Looks like we're going for a kind of 'Partridge Family' thing[28].
  b. "I hate to shiver", says merchant Cathy, "so I found a rugby weight turtle that's as cozy as a cup of cocoa.[29]"
  c. Soaring island real estate triggers reassessment[30].

(7) d. To be very rich or very lovely or very bad or very sure or very good is inherently interesting, since veryness of any kind is not part of dailyness[31].
e. He really knows food but he's not a snootie foodie[32].
f. SQL lets you combine as many expressions as needed to form your **where** clause[33].

3. CONCLUSION. As Example 4 clearly demonstrates, English expression creation is very flexible. The flexibility comes from superposition. None of the expressions in the examples in this paper are self-conscious. Superposition is not something that is learned. Superposition describes a real, underlying mechanism in speech. In some contexts, e.g., (7)c, superposition is so common that the expression may not be able to be stated otherwise or with such efficiency.

This paper has described the two broad categories of superposition in English speech. However, I have not accounted for the role that gestures, other verbal cues, or contexts other than traditional linguistic contexts might play in superposition. It is possible that there are other categories of superposition besides those mentioned here.

---

[1] 'Function' is usually taken to mean syntactic or grammatical function. However, recent work suggests that 'function' might also have semantic aspects.

[2] For the purposes of this paper, a *linguistic expression* means one or more words in a context.

[3] AUG is a functional theory of grammar. Word or expression types are considered to be functions. For example, what most linguists call a 'noun' is, in AUG, a term, a 'verb' is a predicate, and an 'adjective' is a term modifier.

[4] There is almost linguistic Darwinism here. Words or expressions that are 'fit' are repeated and become part of the permanent lexicon. Words or expressions that are not 'fit' are not repeated and, therefore, do not survive.

[5] When I mention 'word meaning' in this paper, I am not suggesting that words are the carriers of meaning. This position has been discredited. I merely wish to use a shorthand expression to describe the semantic properties or processes generated when a word or expression is uttered.

[6] Prasithrathsint, personal correspondence.

[7] It is theoretically possible that a language *L* exists such that *L* does not use any superposition. While I have not done an exhaustive survey of the superposition use of all languages (I don't know how one could, given the linguistic problems involved and the duration of the human life span), I would be greatly surprised to find one.

[8] I will not define *word* in this paper. The reader may use the broadest definition of *word*. The reader should understand, however, that the definition of *word* has a number of implications for the theory of superposition and AUG in general. Those implications will not be discussed in this paper.

[9] A material superposer changes the physical characteristic of a word. The superposer operates upon the word, which is its superponend, to produce a resultant, which is a new word or expression. A superposer is always a sign. AUG is a semiotic theory of language.

[10] I will not discuss whether superposition occurs when only the semantics of a word or expression are modified as, for example, in the various senses of *Oh, that's really nice* caused by stress or intonation.

[11] Depending on one's definition of *word*, one may say that *change* is also contextually superposed if one thinks that the primary function of *change* is predicate.

[12] The *Zoo Run* was the name of a jogging event used to raise money for the Philadelphia Zoo.

[13] I will not discuss the thorny question of whether context is a sign.

[14] In AUG, if an expression functions as a sentence, the expression has type **s**.

[15] To say this another way, words have no inherent type. They get their type, i.e., their function, from the context that they are in. The notions of primary, secondary, tertiary, etc. types is a concession to the obvious fact that, through use, we become familiar with a word W performing a function F more often than it performs the function F'.

[16] Expressions and sentences get their types through a process of reduction. Reduction will not be discussed further in this paper.

[17] AUG claims that languages exist on two levels: the genotype and the phenotype. The genotype level is the most abstract level of language. The genotype level is the same for all languages. The phenotype level is the language-specific expression of the genotype. Superposition occurs on the genotype level of language.

[18] AUG describes one-, two-, and three-place predicates, having the notation Ots, OtOts, and OtOtOts, respectively. In this paper, I use *predicate* to mean any of the three types just mentioned.

[19] AUG does not recognize noun phrases, etc, *per se*. A noun phrase is just another term. In the discussion, I use *noun phrase* to distinguish between a term that is composed of more than one word and a term that consists of one word only.

[20] The complete context or nucleus::satellite relation is *The zoo run*. The word *the* is nonsuperposible and creates the expectation that what follows it has the type term.

[21] Terms are so easy to superpose into term modifiers that it is often difficult to say whether a term modifier is superposed to a term or a term is superposed to a term modifier.

[22] *Antihistamine* may, perhaps, be thought of as a type of medication that causes an action (or prevents an action). However, it does not refer to the action itself.

[23] I think that the pause referred to is an important linguistic fact. However, I will not discuss time in the study of language in this paper.

[24] The study presented here is preliminary. I will not describe my methodology or corpus in detail. The work is continuing.

[25] These last sentences should not be read to suggest that all terms can easily be superposed into term modifiers or predicates. A simple example is the word desk. Some words do not seem to have more than one type, e.g., *whopping*, or are used so frequently to perform their primary function that we may consider them to be nonsuperposible. (Sypniewski 2000)

[26] Repetition turns a temporary expression into a 'permanent' one, i.e., an enduring part of the lexicon.

[27] Block busting in the context of illegal, discriminatory activity means something entirely different. This clearly demonstrates the role of superposition in semantic context.

[28] MANKIEWICZ, JOHN. November 15, 1999. Stock Market Follies. *New Yorker*.

[29] Land's End catalogue, Winter, 1999

[30] headline, *Cape May County Herald*. October 27, 1999.

[31] Gopnick, Adam. August 9, 1999, 'The Stuff of Fame', Comments. *New Yorker*.

[32] Schellinger, L. November 15, 1999. Ming is Rising. *New Yorker*.

[33] Medinets, David. 2000. *PHP3 programming browser-based applications.* New York: McGraw-Hill. (page 134)

REFERENCES

Prasithrathsint, Amara. 1997. The emergence and development of abstract nominalization in Standard Thai. In *Southeast Asian linguistic studies in honor of Vichin Panupong*, ed. by Arthur Abramson. Bangkok: Chulalongkorn University Press.

———. 1994. Borrowing and nominalization of technical terms in Standard Thai. *In Language reform—History and future*, vol. 4. Hamburg: Helmut Buske Verlag.

Shaumyan, Sebastian. 1987. *A semiotic theory of language*. Indianapolis: Indiana University Press.

Sypniewski, Bernard. 1997. Functional superposition. lacus *forum* 23:279–87

———. 2000. Lexical and grammatical superposition in English. lacus *forum* 26:109–16.

Waldbauer, Gilbert. 1996. *Insects throughout the seasons*. Cambridge ma: Harvard University Press.

~

ITERATION AND GENRE IN A 19TH-CENTURY NOVEL

Michael Cummings
York University, Toronto

Among many celebrated minor Dickens characters is Mr. Micawber of *David Copperfield*, appreciated by the general public as a man for whom something must assuredly turn up, although he seems forever destined for further disappointment. He is less generally remembered as a stylist, although the following excerpt from a personal letter so distinguishes him:

(1) ‘my dear sir,
‘Circumstances beyond my individual control have, for a considerable lapse of time, effected a severance of that intimacy which, in the limited opportunities conceded to me in the midst of my professional duties, of contemplating the scenes and events of the past, tinged by the prismatic hues of memory, has ever afforded me, as it ever must continue to afford, gratifying emotions of no common description. This fact, my dear sir, combined with the distinguished elevation to which your talents have raised you, deters me from presuming to aspire to the liberty of addressing the companion of my youth, by the familiar appellation of Copperfield! It is sufficient to know that the name to which I do myself the honour to refer, will ever be treasured among the muniments of our house (I allude to the archives connected with our former lodgers, preserved by Mrs. Micawber), with sentiments of personal esteem amounting to affection. (Project Gutenberg 1996a: Ch. 49)

Both lexical and syntactical aspects of this extraordinary style have a lot to do with Latinism, but there is also a distinguishing textual element, which is a very low incidence of anaphoric co-reference and co-classification (Halliday & Hasan 1985: 73–75), that is, little referential iteration, i.e., iteration of the same semantic referent. There is little iteration of the same lexical item, and there is little semantic iteration with proforms, beyond a number of 1st-person references to the writer himself. This aspect of style could usefully be contrasted with the opposite characteristic in opposite personalities within the same work:

(2) ‘I’ll tell you what I am,’ whispered Mr. Creakle, letting it go at last, with a screw at parting that brought the water into my eyes. ‘I’m a Tartar.’
‘A Tartar,’ said the man with the wooden leg.

> 'When I say I'll do a thing, I do it,' said Mr. Creakle; 'and when I say I will have a thing done, I will have it done.'
>
> '—Will have a thing done, I will have it done,' repeated the man with the wooden leg.
>
> 'I am a determined character,' said Mr. Creakle. 'That's what I am. I do my duty. That's what I do. My flesh and blood'—he looked at Mrs. Creakle as he said this—'when it rises against me, is not my flesh and blood. I discard it. Has that fellow'—to the man with the wooden leg—'been here again?'
>
> 'No,' was the answer.
>
> 'No,' said Mr. Creakle. 'He knows better. He knows me. Let him keep away. I say let him keep away,' said Mr. Creakle, striking his hand upon the table, and looking at Mrs. Creakle, 'for he knows me. Now you have begun to know me too, my young friend, and you may go. Take him away.'
>
> (Project Gutenberg 1996a: Ch. 6)

Mr. Creakle and the man with the wooden leg both suffer from an extreme fondness for referential iteration, which they make to be a sinister symbol of oppression and implacability.

The kind of referential iteration which concerns this paper is the iteration of a reference first introduced in lexical form—thus 'lexical iteration', which may occur in the form of recurring lexis, or as a pro-form. Micawber speech and prose has a very low iterative density, and Creakle speech an extraordinarily high one. Iterative density, in similar fashion to lexical density (Halliday 1994:350–52), can be quantified as the number of iterations through a designated stretch of text. Different parts of the same text can be shown to have different iterative densities by comparing the number of lexical iterations within stretches of similar length.

1. A METHOD FOR MEASURING ITERATIVE DENSITY. In a previous paper, I have described a computer methodology for measuring the change in iterative density continuously through a long text like *Copperfield* (Cummings 1998). The PROLOG language software reads the input text character by character and reconstitutes words as PROLOG lists (Clocksin & Mellish 1984:41ff.) by taking spaces and punctuation as boundary markers. Word lists for lexical words are accumulated in a span list of a size set by the operator at run time. Each such word is parsed to distinguish lexical morphemes from grammatical, and each successive lexical item is compared with the previous items in the span to determine any iteration. Additionally, various pro-forms are recognized as such and reckoned also to be referential iteration. All words are output again as continuous text, with original spacing and punctuation; but each referential iteration is marked with asterisk. At the same time the density of iterations is measured by counting the number of iterations recognized within each successive text segment containing 100 formal items which are potentially iterative, i.e., lexical and pro-form. The end of each such segment is marked in the output text with the serial number of the segment and a numerical figure for the iterative density of that segment.

This method has limitations which are reflected in the output results. The computer doesn’t know that lexical items are not always realized as single words. Neither does it recognize ellipsis, the distinction between pronominal and dummy subject ‘it’, or that between deictic and dummy subject ‘there’. The parsing of lexical and grammatical morphemes depends on a very limited algorithm which doesn’t always correctly manage the plural of nouns and some other sorts of morphotactics. The segments of text measured for density are exclusive, not continuously overlapping, so that the segmentation of the text is a kind of sampling.

Although the main subject of this paper is not *Copperfield*, but a much shorter and therefore more manageable Dickens novel, *Oliver Twist* (Project Gutenberg 1996b), the computer results can be illustrated with an output segment from the processing of *Copperfield* which constitutes the analysis of the Creakle passage (with the lexical word iteration span set for 30; Cummings 1998:258 and nt. 2):

(3) ‘*i’ll [180-48] tell you what *i am,’ whispered mr. *creakle, letting *it go at
last, with a screw at parting that brought the water into *my eyes.
‘*i’*m a tartar.’
‘a *tartar,’ said the man with the wooden leg.
‘when *i say *i’*ll do a thing, *i do *it,’ said mr. *creakle; ‘and when
*i *say *i will have a *thing done, *i will have *it *done.’
‘- will have a *thing *done, *i will have *it *done,’ repeated the *man
with the *wooden *leg.
‘*i am a determined character,’ said mr. *creakle. ‘that’s what *i
am. *i do my duty. that’s what *i do. *my flesh and blood’ - *he
looked at mrs. *creakle as *he said this - ‘when *it rises against *me,
is not *my *flesh and *blood. *i discard *it. has that fellow’ - to
the *man with the *wooden *leg -’been here again?’
‘no,’ was the answer.
‘no,’ said mr. *creakle. ‘*he knows better. *he *knows *me. let *him
keep away. *i say *let *him *keep away,’ said mr. *creakle, striking
*his hand upon the table, and looking at mrs. *creakle, ‘for *he *knows
[181-62] *me. now you have begun to *know *me too, *my young friend,
and *you

This selection from the output includes all of segment #181, which has a density of 62 recognized iterations, each of which is preceded by an asterisk. By comparison, the segment which contains most of the above quotation from Mr. Micawber’s letter (#1554), contains only 23 recognized iterations, 14 of them within the text quoted. Examination of the output segment reveals the effect of some of the limitations in the computer method. The computer has been instructed that ‘am’ and ‘will’ are to be regarded as function words, not lexical items, but has not been instructed that “‘m” and “‘ll” are variants whose repetition should not be marked. The computer has been instructed that ‘do’ is more probably a function word, and so misses the iterations of

I: 1-25 2-15 3-15 4-34 5-32 6-20 7-27 8-26 9-18 10-19 11-16 12-38 13-38 14-35 15-30 16-37 17-25 . . . 862-23 863-25 864-23 865-21 866-17 867-24 868-24

II: 8-321 8-794 11-322 11-814 12-329 12-594 13-45 13-173 13-320 13-599 13-795 13-854 14-215 . . . 43-66 43-95 43-210 44-636 44-718 45-161 45-731 47-644 47-749 48-807 49-377

III: 32-320 36-319 40-793 42-794 44-513 47-597 49-249 51-723 52-70 52-388 52-693 53-9 53-172 . . . 116-642 116-705 116-747 118-635 118-704 119-730 119-748

Figure 1. *Condensed logs of segment densities in Oliver Twist.*

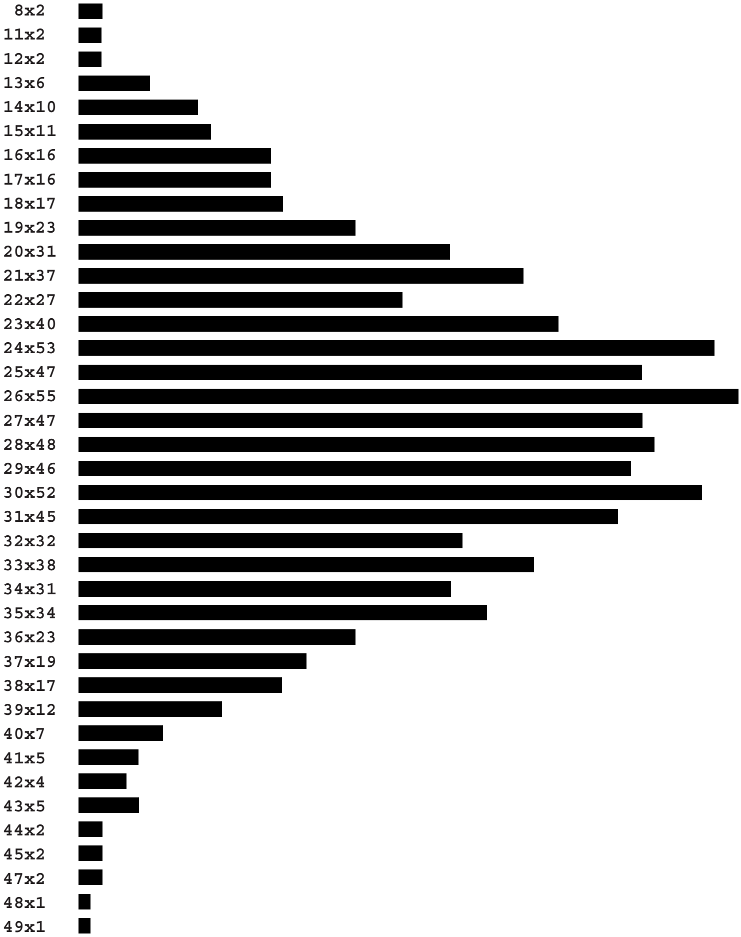


Figure 2. *Frequency distribution graph for segment densities in* Oliver Twist.

lexical 'do' in this passage. The computer has also been told that 'that' is more likely a modifier than a pro-form, and so it misses the use of 'that' when it is a pro-form. The results are therefore heuristic rather than absolute; in this case the computer method has actually underestimated the degree of referential iteration.

The purpose of this paper is to show how with the help of the computer, however imperfect, the pattern of iterative densities in a long text can be correlated with other stylistic variables, specifically genre variables (Eggins 1994: 25–48; Martin 1992:546ff.). Another kind of output produced by the computer is very useful for this purpose: a log of the densities in each segment. This takes three forms, which are illustrated in Figure 1, for *Oliver Twist*. The first type of log is a list of all the densities recorded for the 868 segments of *Twist* in serial order. The second type is the same list ordered from the lowest densities to the highest, with the order of serial number and density figure reversed. The third type is like the second, but each pair of figures represents an accumulated density for three consecutive segments ending with the segment designated by the serial number.

Yet another computer output is represented in Figure 2, a frequency distribution graph for the various densities. In this graph, each density, from the lowest to the highest, and the number of times it occurs are shown in the left-hand column. The purpose of these secondary outputs is to enable the identification and comparison of segments with similar degrees of referential iteration, with a view to making stylistic statements about genre variables instantiated in such segments.

2. ITERATION AND GENRE. In order to make stylistic statements about segments of text from *Oliver Twist* whose iterative densities have been measured, it is necessary to characterize these segments as belonging to one genre or another. The whole text of *Twist* can be described in terms of just five elementary genre categories, *narrative, descriptive, expository, monologue,* and *dialogue*. The most elementary features of these categories can be illustrated in examples of no more than two or three consecutive sentences each. Narrative, as in example (4), has the feature sequence of events in time. Descriptive, as in example (5), lacks sequence of events in time, but shares with narrative the feature visualizable scene. Example (6) represents expository genre, with the feature abstract facts and evaluations (on the part of the novelistic narrator). Example (7) represents monologue, being an extended quoted address by one fictional character. Dialogue, as in example (8), has the feature quoted address by two or more characters to one another.

(4) **Narrative**

as *they approached the city, the noise and traffic gradually increased; when *they threaded the streets between shoreditch and smithfield, *it had swelled into a roar of sound and bustle.

(5) **Descriptive**
*it was market-morning. the ground was covered, nearly ankle-deep, with filth and mire; a thick steam, perpetually rising from the reeking bodies of the cattle, and mingling with the fog, *which seemd to rest upon the chimney-tops, hung heavily above. all the pens in the centre of the large area, and as many temporary *pens as could be crowded into the vacant space, were filled with sheep; tied up to posts by the gutter side were long lines of beasts and oxen, three or four deep.

(6) **Expository**
i need [172-20] hardly beg the reader to observe, that this action should tend to exalt *them in the opinion of all public and patriotic men, in almost as great a degree as this strong proof of *their anxiety for *their own preservation and safety goes to corroborate and confirm the little code of laws *which certain profound and sound-judging philosophers have laid down as the main-springs of all nature's deeds and actions . . .

(7) **Monologue**
'bill, dear *bill, *you cannot have the heart to kill *me. oh! think of all *i have given up, only this one *night, for *you. *you shall have time to *think, and save yourself this crime; *i will not loose *my hold, *you *cannot throw *me off.

(8) **Dialogue**
'I mean to say that *i shouldn't,' *replied *tom, angrily. '*there, now. ah! *who'll *say as much as that, *i should like to know; eh, *fagin?'

'nobody, *my *dear,' *replied the *jew; 'not a soul, *tom. *i don't *know one of 'em that would do *it besides you; not one of '*em, *my *dear.'

In order to evaluate the whole output for correlations between iterative density and genre variation, the 868 segments in the output file were reordered in a word processor by density from lowest to highest. The resulting arrangement corresponds to the pattern represented by the second type of log in Figure 1 and in the frequency distribution graph of Figure 2. Then four samples, each of 30 successive segments (32 in Sample 4), were taken from the resulting file. The location of these samples along the density axis is represented in Table 1.

The selection of segments to sample is thus based on a compromise between the distribution of densities and the numbers of segments. In terms of densities, the gap between samples is fairly constant, at 7 or 8. In terms of the number of segments, the first and last samples mark the ends of the file, the second and third samples are equally spaced from the ends and from the centre.

After the samples were isolated, each of the segments in the samples was coded for genre type. When a segment contained a boundary between two genres, the segment

Sample	Segments	No. of Segments	Range of Densities
Sample 1	1-30	30	8-15
not in samples	31-209	179	15-23
Sample 2	210-239	30	23
not in samples	240-627	388	23-31
Sample 3	628-657	30	31-32
not in samples	658-836	179	32-39
Sample 4	837-868	32	39-49

***Table 1**. Sampling of segments measured for iterative density.*

was labelled with the genre name that represented a clear preponderance of text. If there was no clear preponderance, then the segment was labelled mixed. Additionally, segments which did not belong to dialogue or mixed genre were labelled as either *concentrated* or *diffuse*. A segment is concentrated if it focuses on just one or two concrete (i.e., not abstract) participants, and diffuse if it focuses on more than two. The formal criterion for focussing on concrete participants is the occurrence of such participants as thematic Subject elements in independent clauses. (A thematic Subject is a grammatical Subject which is not inverted to a post-main-verb position; cf. Berry 1996.) Dialogue genre was reckoned to focus inherently on two concrete participants, the parties to the dialogue, and thus to be inherently concentrated.

The data indicate that segments of whatever genre coded as diffuse tend to have low iterative density, and segments regarded as concentrated tend to high density. Good illustrations are provided by examples (9) and (10), diffuse descriptive and concentrated monologue segments respectively. Example (9) with an iterative density of 8 has both narrative and descriptive elements, with descriptive clearly prevailing. It is diffuse insofar as thematic Subjects of independent clauses include 'Mr. Sikes', 'The ground', 'a thick steam' and several others. Example (10) with density of 41 has narrative, monologue, and dialogue elements, with monologue prevailing. It is concentrated in the sense that thematic Subjects of independent non-dialogue clauses include only Nancy and Sikes.

(9) **Description, diffuse**

square, mr. sikes struck, by way of chiswell [320-13] *street, into barbican: thence into long lane, and so into smithfield; from *which latter place arose a tumult of discordant sounds that filled oliver twist with amazement.

*it was market-morning. the ground was covered, nearly ankle-deep, with filth and mire; a thick steam, perpetually rising from the reeking bodies of the cattle, and mingling with the fog, *which seemd to rest upon the chimney-tops, hung heavily above. all the pens in the centre of the large area, and as many temporary *pens as could be crowded into the vacant space, were filled with sheep; tied up to posts by the gutter side were long lines of beasts and oxen, three or four deep. countrymen, butchers, drovers, hawkers, boys, thieves, idlers, and vagabonds of every low grade, were mingled together in a mass; the whistling of *drovers, the barking dogs, the bellowing and plunging of the *oxen, the bleating of sheep, the grunting and squeaking of pigs, the cries of *hawkers, the shouts, oaths, and quarrelling on all sides; the ringing of bells and roar of voices, that issued from every public-house; [321-8] the crowding,

(10) **Monologue, concentrated**

mortal fear,--'i- [752-28] -*i won't scream or cry--not once--hear me--speak to *me--tell *me what *i have done!'

'you know, *you *she devil!' returned the robber, suppressing *his breath. '*you were watched to-night; every word *you said was heard.'

'then spare *my life for the love of heaven, as *i *spared yours,' rejoined the girl, clinging to *him. 'bill, dear *bill, *you cannot have the heart to kill *me. oh! think of all *i have given up, only this one *night, for *you. *you shall have time to *think, and save yourself this crime; *i will not loose *my hold, *you *cannot throw *me off. *bill, *bill, for *dear god's sake, for your own, for mine, stop before *you spill *my blood! *i have been true to *you, upon *my guilty soul *i have!'

the man struggled violently, to release *his arms; but those of the girl were clasped round *his, and tear *her as *he would, *he could not *tear [753-41] *them away.

That diffuse tends to low density and concentrated to high density is indicated in Table 2, which represents the distribution of genre types and concentration of focus across all the samples. The number of diffuse segments drops off rapidly after Sample 1. The number of segments labelled as concentrated increases slightly through Samples 1 - 3; but the combined number of segments labelled concentrated and dialogue segments grows rapidly across the samples. Other genre correlations revealed by Table 2 are the tendency for narrative and descriptive genre to show low iterative density, whereas monologue and especially dialogue show a preference for high iterative density.

3. TYPES OF NARRATIVE. Not surprisingly, Table 2 indicates that, after dialogue genre, the vast majority of unmixed segments are the narrative ones. Very characteristic of narrative segments with the lowest iterative densities (Sample 1) is a subtype which depicts a continuously advancing point of view, as on a journey. In example (11) the point of view belongs to Sikes and Oliver on a foot journey into the heart of London. The movement here is particularly complex because along two axes simultaneously, time and space; the new phenomena encountered increase in variety and number along both, and naturally make for a very limited number of iterations.

Genre:	Sample 1		Sample 2		Sample 3		Sample 4		TOTALS	
	No.	%	No.	%	No.	%	No.	%	No.	%
Narrative	19	15.6	15	12.3	8	6.6			42	34.4
Descriptive	5	4.1			1	0.8			6	4.9
Expository	1	0.8							1	0.8
Monologic	1	0.8	2	1.6	2	1.6	5	4.1	10	8.2
Dialogic			6	4.9	15	12.3	24	19.7	45	36.9
Mixed	4	3.3	7	5.7	4	3.3	3	2.5	18	14.8
TOTALS	30	24.6	30	24.6	30	24.6	32	26.2	122	100
Concentrat.	4	3.3	5	4.1	6	4.9	1	0.8	16	13.1
Diffuse	22	18.0	12	9.8	5	4.1	4	3.3	43	35.2
TOTALS	26	21.3	17	13.9	11	9.0	5	4.1	59	48.3

Table 2. *Distribution of genre types*

(11) **Narrative, diffuse.**

> wrong side of the road, [319-15] had endangered *his arriving at the office, a quarter of a minute after *his time. the public-houses, with gas-lights burning inside, were already open. by degrees, other shops began to be unclosed, and a few scattered people were met with. then, came straggling groups of labourers going to *their work; then, men and women with fish-baskets on *their heads; donkey-carts laden with vegetables; chaise-*carts filled with live-stock or whole carcasses of meat; milk-*women with pails; an unbroken concourse of people, trudging out with various supplies to the eastern suburbs of the town. as *they approached the city, the noise and traffic gradually increased; when *they threaded the streets between shoreditch and smithfield, *it had swelled into a roar of sound and bustle. *it was as light as *it was likely to be, till night came on again, and the busy morning of half the london population had begun.
>
> turning down sun *street and crown *street, and crossing finsbury square, mr. sikes struck, by way of chiswell [320-13] *street, into

Narrative segments in the next band of iterative densities (Sample 2) show less of this subtype, proportionally more of a more static kind of narrative. In example (12) the point of view starts with Oliver, moves to the things he looks at, then moves back to and remains with Oliver. The limiting of attention to a single character and his immediate surroundings makes for repeated mention of the same limited set of participants, and thus for a higher iterative density.

(12) **Narrative, diffuse.**

> candle was burning low in the socket when *he [99-23] rose to *his feet. having gazed cautiously round *him, and listened intently, *he gently undid the fastenings of the door, and looked abroad.

> *it was a cold, dark night. the stars seemed, to the boy's eyes, farther from the earth than *he had ever seen *them before; *there was no wind; and the sombre shadows thrown by the trees upon the ground, *looked sepulchral and death-like, from being so still. *he softly reclosed the *door. having availed *himself of the expiring light of the candle to tie up in a handkerchief the few articles of wearing apparel *he had, sat *himself down upon a bench, to wait for morning.
>
> with the first ray of *light that struggled through the crevices in the shutters, oliver arose, and again unbarred the *door. one timid look around--one moment's pause of hesitation--*he had closed *it behind *him, and was in the open street.
>
> *he *looked to the right and to the left, uncertain whither to fly.
>
> *he remembered to have seen the waggons, as *they went out, toiling up the hill. *he took the same route; [100-23] and arriving at a footpath

A third and even more intriguing narrative style approaches the level of iterative density characteristic of dialogue genre. This is a sequence of events describing characters who communicate with one another in an unspoken dialogue. In example (13) Nancy, Sikes, Fagin, and even the dog communicate among themselves, and even to themselves, by gesture and by the reported event of dialogue. This makes for a great many iterative references to the same set of participants, as in dialogue proper.

(13) **Narrative, diffuse.**

> and a half closing of *his deeply- [222-23] set *eyes, warned miss *nancy that *she was *disposed to be too communicative, is not a matter of much importance. the fact is all we need care for here; and the *fact is, that *she suddenly checked *herself, and with several gracious smiles upon mr. sikes, turned the conversation to other *matters. in about ten minutes' time, mr. fagin was seized with a fit of coughing; upon *which *nancy pulled *her shawl over *her shoulders, and declared *it was *time to go. mr. *sikes, finding that *he was walking a short part of *her way *himself, expressed *his intention of accompanying *her; *they went away together, followed, at a little distant, by the dog, *who slunk out of a back-yard as soon as *his master was out of sight.
>
> the jew thrust *his head out of the room door when sikes had left *it; looked after *him as we walked up the dark passage; shook *his clenched fist; muttered a deep curse; and then, with a horrible grin, reseated *himself at the table; *where *he was soon *deeply absorbed [223-31] in the interesting pages of the hue-and-cry.

4. A NOTE ON THE NARRATOR. The fact of only one unmixed expository segment should be noted. The narrator rarely intrudes with a lengthy general philosophizing in this novel, thus affording only one clear segment of the genre in the samples.

However the quotation from it in example (6) could remind its reader of another voice earlier alluded to—Mr. Micawber! Ironically, the prose of Micawber and the prose of Dickens as expositor here share a significant characteristic—a low iterative density. In fact the density of the expositor's segment in *Twist* (13) is much lower than that of the Micawber segment from *Copperfield* (23), and probably owes this trait to the same facetiousness in choice of lexis, which the reader is expected to judge ironic in the case of the general narrator, but an eccentricity in the case of Micawber.

REFERENCES

BERRY, MARGARET. 1996. What is theme?—A(nother) personal view. In *Meaning and form: Systemic functional interpretations*, ed. by Margaret Berry, Christopher Butler, Robin Fawcett & Guowen Huang, 1–64. Norwood NJ: Ablex Publishing.

CLOCKSIN, W. F. & C. S. MELLISH. 1984. *Programming in PROLOG*, 2nd ed. New York: Springer Verlag.

CUMMINGS, MICHAEL. 1998. Lexical repetition and the given/new distinction in written English *LACUS forum* 24:253–66.

EGGINS, SUZANNE. 1994. *An introduction to systemic functional linguistics*. London: Pinter.

HALLIDAY, M.A.K. 1994. *An introduction to functional grammar*, 2nd ed. London: Edward Arnold.

——— & RUQAIYA HASAN. 1985. *Language, context and text: Aspects of language in a social-semiotic perspective*. Geelong, Vic.: Deakin University Press [also Oxford University Press, 1989].

MARTIN, JAMES. 1992. *English text: System and structure*. Amsterdam: Benjamins.

Project Gutenberg. 1996a. *The Project Gutenberg Etext of David Copperfield by Charles Dickens.* [Etext #766]. <http://promo.net/pg/index.html>

Project Gutenberg. 1996b. *The Project Gutenberg Etext of Oliver Twist by Charles Dickens.* [Etext#736]. <http://promo.net/pg/index.html>

~

VERBAL REASONING AND BRAIN HEMISPHERES

Inga B. Dolinina
McMaster University

INTRODUCTION. In this paper I discuss the results of a study of the role of the left and the right hemispheres in the verbal reasoning task of solving syllogisms. This study aimed to shed neurological light on whether educated individuals who have developed an ability to reason logically still preserve and use their pre-logical way of reasoning. It therefore investigated whether and how these two modes of reasoning—'empirical' (pre-logical, non-formal) and 'theoretical' (formal, logical)—are represented neurologically. I argue that the results of this study require us to reconsider the accepted position that only the right hemisphere processes information holistically and to recognize that the left hemisphere is responsible for the recognition of inter-propositional linguistic wholes.

1. PRIOR STUDIES IN REASONING MODES. To give an idea of the reasoning modes I cite an example of typical empirical and theoretical answers to syllogisms observed by Scribner (1977:485).

The syllogism:

(1) All people who own houses pay house tax. Boima does not pay a house tax. Does he own a house?

can have in actual discourse two types of answers: a formally correct one, exemplifying a theoretical mode (signalled by the letter 't'):

(1t) No, he does not,

and a formally flawed one, exemplifying an empirical mode (signalled by the letter 'e'):

(1e) Yes, he has a house. But he does not pay the tax, because he has no money.

Though the first answer is logically correct, the second 'incorrect' one may be true to reality.

A variety of disciplines—cross-cultural, developmental, educational—and scholars—psychologists, philosophers, semioticians—have addressed the existence of these two modes of reasoning (Dolinina in press).

Cross-cultural observations (Lévy-Bruhl 1923, Lévi-Strauss 1966), confirmed by cross-cultural experiments (Luria 1976/30's; Cole, Gay, Glick & Sharp 1971; Cole & Scribner 1974; Scribner 1977; etc.), discovered that individuals in a 'traditional' society can use only an empirical ('pre-logical') mode, whereas individuals in 'advanced' societies can use a theoretical ('logical') mode.

Two different explanations were offered as to why individuals from traditional societies do not develop a theoretical mode of reasoning. 'Literacy theory' connected appearance of the theoretical mode with literacy (Olson et al. 1985; Ong 1982). This explanation proved to be incorrect, since the theoretical mode is not used by traditional cultures with local literacy—the case of the Vai discussed by Scribner and Cole (1981). Evidently, only western-type literacies, which go back to the Greek tradition of reasoning, have this effect on cognitive development.

The 'discourse theory' explanation connected inability to use a theoretical mode with the absence in the traditional culture of such a 'script' as a syllogism. Discourse theory argues that the semantic resolution of a syllogistic script is fully dependent on its inner content and on the rules for relating the content of the premises with one another. It does not project the content against the realities of the outside world, which is actually the basis of the empirical strategy (see discussion in Scribner 1977, Olson et al. 1985, Ong 1982). Unlike 'literacy' theory, the discourse theory explanation looks highly plausible. But if it is correct it raises another question: do individuals who acquire the ability to work within a syllogistic script, that is within a theoretical mode, a) completely switch from empirical to formal reasoning, b) use both strategies (that is, manifest 'thought heterogeneity'), or c) still stick to the empirical mode?

On this issue, cross-cultural and semiotic researchers (e.g., Lotman's School of Semiotics in Tartu University—Lotman 1978, Ivanov 1978, etc.) suppose that both reasoning strategies are utilized, but psychologists are not so unanimous.

Thus, Piaget's developmental research (Piaget 1964, 1971, 1975; Inhelder & Piaget 1951) discusses the formation of reasoning abilities. Piaget claims that formal reasoning appears quite late (between 11–15), in the last, fourth, stage of cognitive development, and requires formation not only of intra-propositional connections (between the predicate and its nominals) but also of inter-propositional connections (between clauses: causal, inferential, etc.). In his early works Piaget claims that formal reasoning always develops, but in formulating his later theory of equilibration he comments that its development is socially triggered. But he does not specify whether the fourth stage always emerges, or what happens to empirical reasoning when formal reasoning is established.

Contemporary cognitive psychologists are particularly interested in the 'fate' of the theoretical mode once it is established (Wason & Johnson-Laird 1972; Johnson-Laird 1983; Johnson-Laird & Byrne 1991; Moore 1986; Wilson & Sperber 1986). They raise two questions: 1) do individuals who acquire the theoretical mode actually use it instead of the empirical mode? and 2) is a theoretical mode, that is, formal reasoning, represented in the mind as a special mental component? Some cognitive psychologists recognize the dominance of a theoretical mode once it is established and claim that it is represented in the brain (e.g., Wilson & Sperber 1986), but others argue that people continue to use empirical modes of reasoning in normal-life situations and that the theoretical mode is not a special mental component.

Thus, different approaches gave diverse answers as to the status of the two modes in human discourse and in the human brain. The neurological experiments discussed below aimed to answer this question from a new perspective. Considering the consensus that contemporary educated adults have potential access to both modes, they aimed to find out whether use and selection of a mode is somehow determined 1) by lateralized functions of the brain hemispheres or 2) by the type of information contained in the syllogism.

2. DESIGN OF THE NEUROLOGICAL RESEARCH ON REASONING MODE PREFERENCES. The study was carried out in the Sechenov Institute of Evolutionary Physiology and Biochemistry, Russian Academy of Sciences, St. Petersburg, Russia, within the project of Functional Specialization of Brain Hemispheres. The data were collected from psychiatric patients (predominantly with depression) while they were recovering from transitory suppression of one of the hemispheres by electroconvulsive therapy. Special attention was paid to contributions of the right hemisphere to language, a most controversial issue even now. The project started in the early 1970s. The first report (Balonov & Deglin 1976) described the perception and production of sounds under temporary suppression by unilateral electroshocks of one hemisphere, leaving the other practically normal. Further research investigated syntax, morphology, lexicon, idioms and metaphors, and lastly syllogisms. The results were reported in a variety of publications in Russian, but very few in English and German and not in linguistic journals; the only linguistic response I am aware of is Jakobson 1995/1980.

The research on syllogisms aimed 1) to test the abilities of literate adults to use both reasoning modes and 2) to discover the contribution of the left and right hemispheres to solving syllogisms. For this purpose each subject's performance was tested in three states: before electroshock, and after transitory suppression of the left or right brain hemispheres (Černigovskaja & Deglin 1990, Deglin 1995, Deglin & Kinsbourne 1996). The question arises how subjects with a suppressed left hemisphere (including Wernicke's area responsible for language processing) are able to handle syllogisms at all. The answer is twofold. First, the restitution of hemispheric functions after unilateral electroshocks, though rather quick (30–40 minutes), is gradual. This fact, and the results, suggest that subjects with left hemisphere suppressed were processing syllogisms when the left hemisphere had partially restored its functions and was able to process intraclausal connections, though not yet the interclausal ones needed for formal reasoning. Second, the studies of this group demonstrated that the right hemisphere has considerable involvement in language processing—it evaluates the semantic validity of linguistic information and verifies it against reality. So the right hemisphere may independently process the content of the propositions which constitute syllogisms. Thus, with minimal restoration of the syntactic function by the left hemisphere (considering the simplicity of the syntactic structures involved) the individuals are able not only to understand, but also to produce, sentences, and consequently 'quasi-syllogisms', though at this stage they cannot process an inter-propositional complex which forms the basis of a real syllogism.

The tested group included 14 (for Experiment 1) and 10 (for Experiment 2) right-handed individuals of both sexes, all with secondary and some with university education. The study tested solving of two types of syllogisms:

a) Experiment 1: syllogisms with true premises (including syllogisms with both familiar and unfamiliar content similar to the ones in Luria's experiments)
b) Experiment 2: syllogisms with a false premise (a type of syllogisms was never tested before).

3. RESULTS OF EXPERIMENT 1: SOLVING SYLLOGISMS WITH TRUE PREMISES. Here are some examples of syllogisms offered to the subjects:

(2) There are fish in all the rivers where people put nets. People put nets on the Neva River. Are there fish in the Neva River, or not?

(3) Every state has a flag. Zambia is a state. Has Zambia a flag, or not?

Characteristics of answers to the syllogisms from the subjects in each of the three states:

CONTROL STATE (both hemispheres active): In this state subjects gave predominantly theoretical answers (12 of 14). Such a result could be expected, since all the subjects were educated within a culture in which syllogisms are used. Only two subjects sometimes gave empirical responses (in accordance with their experiences and/or beliefs) to some syllogisms, like the following in response to (2), which is classified as empirical because it is based on knowledge of realities, not just on the content of the premises:

(2e) Everybody knows that there are smelt in the Neva.

RIGHT HEMISPHERE SUPPRESSION (left hemisphere active): In this state the same number of individuals (2) gave some empirical answers. But as a whole there were strong differences from the control group.

First, a more pronounced tendency for usage of a theoretical mode was observed.

Second, the linguistic behavior of the subjects changed: it became more spontaneous; all the tasks were solved more readily, without hesitation, and with much more assurance than in the control group.

Third, in justifying their answers, the subjects referred spontaneously to the contents of the premises, and not to their worldly knowledge. Examples of such answers are:

(2t) Since it is said here that in all the rivers where people put nets there are fish, and the Neva is a river, and people put nets on it, so there are.

(3t) Since each state has a flag, so Zambia also has one.

LEFT HEMISPHERE SUPPRESSION (right hemisphere active): In this case a strong difference from both previous cases was observed.

First, the number of empirical answers dramatically increased: 11 subjects (not just 2) out of 14 used them. Some subjects gave only empirical answers without using the theoretical mode at all. In comparison with the control state, where only some syllogisms, usually those like (2) or (3) with strongly familiar or strongly unfamiliar content, were given empirical answers, here all syllogisms independently of the type of content were given typical empirical answers, such as:

(2e') Of course there are, but inedible, they cannot be edible, there is much machine oil. Smelt should be.

(3e) Who knows it, this Zambia? How can I know whether it has a flag or not?

Second, though empirical answers prevailed, there were far fewer empirical answers to syllogisms with unfamiliar content than to syllogisms with familiar content. In this respect these educated subjects behaved quite differently from Luria's non-schooled subjects, who refused even to discuss syllogisms with unfamiliar content, claiming that they cannot judge about anything they do not know.

Third, the subjects' readiness to use the modes differed from the control group, and was the reverse of the case of a suppressed right hemisphere: empirical answers were given quickly and with full assurance, whereas theoretical answers were given with difficulty and hesitations.

Experiment 1 thus suggested that one and the same person who has access to both reasoning modes may solve one and the same task differently in different states. The type of answer depends mainly on which hemisphere is active, and also to some extent on the familiarity of the content of the premises. But overwhelmingly the left hemisphere uses a theoretical mode, the right hemisphere an empirical mode. So this experiment showed 'that within our culture, under usual conditions the right-hemisphere mode of thought [empirical mode - I.D.] is not drawn to syllogism solving' (Deglin 1995: 23–24). Thus, in respect to syllogistic reasoning there is a strong lateralization of the functions of the hemispheres: this mode of reasoning is carried on by the left hemisphere.

4. RESULTS OF EXPERIMENT 2: SOLVING SYLLOGISMS WITH FALSE PREMISES. This experiment aimed to find out how and whether the hemispheres are involved in identifying the quality of information contained in the syllogisms.

Examples of a syllogism with false premises (marked with*) used in the test:

(4) All monkeys climb trees.*The hedgehog is a monkey. Does the hedgehog climb trees, or not?

(5) *Winter is cold in tropical countries. Ecuador is a tropical country. Is it cold in winter in Ecuador, or not?

Observed responses were of four types: theoretical, empirical mode (two types) and 'ambivalent'. Typical theoretical responses followed the rules of making inferences, relying only on the inside-the-syllogism information, thus showing absence of criticism of the quality or reliability of information presented by the syllogism:

(4t) Yes, it (hedgehog) does.

(5t) Yes, winters are cold in Ecuador.

Empirical responses were of two types. One type rejected the content of a premise:

(4e) The hedgehog is not a monkey,

(5e) It is warm in Ecuador! It is a tropical country!

The other type, similar to that of Luria's non-schooled subjects when they were offered syllogisms with unfamiliar content, represented a complete refusal to answer, justified by the absence of relevant personal experience (no outside-the-syllogism information):

(4e') I have not seen hedgehogs, I do not know whether they climb trees or not.

'Ambivalent' answers (signalled by the letter 'a') exhibit a difficulty in choosing between the modes. This type of answer was met only in the control state, never in a state of suppression of either hemisphere:

(5a) Due to the structure of what is written, the answer should be positive, but I think winters cannot be cold there.

General results of solving syllogisms with a false premise were the following:

THE CONTROL GROUP gave all four types of responses. Predominantly (⅔ of answers) empirical responses were used—rejection of the false premise or refusal to solve the syllogism. About ⅓ of the answers were theoretical (ignoring the premise's mismatch with reality). The few ambivalent responses explicitly verbalized the subjects' hesitancy about which strategy to use—to follow the rules of inference ignoring the false premise, or to pursue the truth:

(4a) Must I answer so as it is written here? Then the hedgehog climbs trees. But it does not climb. It is not a monkey.

LEFT HEMISPHERE SUPPRESSION (right active) was characterized by very strong rejection of false premises, that is, by using an empirical mode (slightly under 90%). The subjects refuted false premises with conviction and with a strong emotional reaction: extreme indignation, extended denials, etc:

(5e') It is warm in Ecuador! Rubbish! It is a tropical country! It can't be winter there, because it is in the South. Wrong what you write!

RIGHT HEMISPHERE SUPPRESSION (left active) produced a dramatic change both from the control state and from the left hemisphere suppression state. The percentage of theoretical answers grew to 75%, whereas that of empirical answers decreased to about 25%. Some individuals did not use empirical answers at all. The discourse behavior of the subjects who gave theoretical answers was strikingly confident: they did not pay any attention to the falsehood of the premises (completely relying on the authority of what is 'said' and 'written'), and proceeded with the syllogism without even thinking about the quality of the information given to them. Besides, the emotional attitude to the task radically changed—the subjects performed calmly, neglecting the absurdity of the premises. The result was a variety of absurd conclusions, derived in accordance with correct rules of formal logic:

(4t) The hedgehog climbs trees.

(5t) It is cold in Ecuador.

5. CONCLUSIONS. This neurological research:

1) showed that both theoretical and empirical modes are used by contemporary individuals and are simultaneously present in the brain of one and the same person. So the experiment confirmed the reality of 'thought heterogeneity'.
2) suggested that both hemispheres participate in reasoning, but that reasoning functions are lateralized: the left hemisphere controls the theoretical mode, the right the empirical mode.
3) established a certain tendency in selection of the modes: the empirical mode is preferred for syllogisms with familiar content, the theoretical mode for syllogisms with unfamiliar content.
4) discovered a special role of the right hemisphere: it controls the quality of information in the syllogism, whereas the left hemisphere is indifferent to the content of the premises.

6. DISCUSSION. Combining these conclusions with the results of prior studies of reasoning modes and of the role of the hemispheres in language processing allows us to discuss 1) the existence of a distinctive neurological substrate in contemporary schooled ('bi-mode') individuals for each mode, 2) the similarities and differences in using reasoning patterns between 'mono-mode' individuals on the one hand and 'bi-mode' individuals with a suppressed left hemisphere on the other, and 3) the 'semantic responsibilities' of the right and left hemispheres.

1) The neurological experiments confirmed the observations of cross-cultural research that educated subjects as a rule use both reasoning strategies, and demonstrated that use of these strategies has a neurological substrate, since each hemisphere operates with a different mode. This, in its turn, suggests that each mode has a mental representation in the relevant hemisphere—the theoretical mode in the left hemisphere, the empirical mode in the right hemisphere. The lateralization of reasoning modes can be regarded as a counterevidence to Johnson-Laird's claim that formal reasoning, unlike semantically based reasoning, is not represented in the mind. If his claim were correct, how could one explain the ability of 'left hemispheric' individuals (under the suppression of the right hemisphere and thus with blocking out the ability to utilize an empirical mode and, consequently, to actualize a semantic-based strategy) to reason, and to use exclusively a theoretical mode with full neglect of outside-the-syllogism realities? To be actualized (even despite semantic incongruencies with the realities of the world, as in syllogisms with false premises), this mode needs to be represented in the mind. Thus the step-by-step 'grammar of inference' must be a self-sufficient mental component. Thus formal rules of inference cannot be reduced to semantically adequate scenarios, as Johnson-Laird's semantic model of reasoning requires.
2) The neurological experiments demonstrate that reasoning carried out by bi-mode individuals with only the right hemisphere active is similar to the reasoning of mono-mode individuals (those from 'traditional' societies and adolescents who have not yet developed the last stage of cognitive development and lack formation of inter-propositional connections) in that both groups reason in an empirical mode. This discovery gives further support to claims of semioticians (already supported by developmental and neurological studies) that the empirical mode (the function of the right hemisphere) is more archaic, and precedes the appearance of the theoretical mode. There is though, as these experiments revealed, a pronounced difference in how both groups deal with information outside of their personal knowledge: mono-modes completely refuse to make any inferences, whereas bi-modes have a tendency to use the theoretical mode (which supports the claims of some developmental models of reasoning that in acquiring new scientific knowledge the theoretical mode is predominantly used).

One more point is important in the comparison of the two groups: their perception of the syllogism. Cross-cultural experiments (Luria's and later) discovered that the mono-modes were unable to repeat the syllogism, because of their inability to perceive it as one semantic unit, as a whole. Though this particular aspect of treating syllogisms was not directly addressed in this neurological experiment, the responses given made it obvious that the bi-modes with suppressed left hemisphere also failed to perceive the syllogism as a whole.

3) Since it is universally accepted that the right hemisphere processes information in a holistic way, and the left hemisphere non-holistically, the inability of the right hemisphere to recognise the syllogism as a whole, and the contrasting ability of the left hemisphere, forces us to modify our understanding of the concept of holicity. The results described here are congruent with the linguistic functions of the right hemisphere discovered in other experiments of the same group (since the early 1970s). These experiments showed that the right hemisphere plays a much more considerable role in processing semantic information than was earlier supposed, and is not restricted to pragmatic aspects or processing of polysemy. This implies that both hemispheres, not just the left, are involved in language processing. According to the 'bi-hemispheric' approach the right hemisphere is responsible linguistically for (among other things) the referential and semantic correctness of words with respect to real-world denotates (including first-order predicates), and the left hemisphere for their syntactic organization (Balonov, Deglin & Dolinina 1983) and also for all kinds of 'connectors', such as subordinators, and also for second-order predicates, especially modal, phasal, etc. verbs. So both hemispheres have semantic responsibilities, but their scope is quite different: if the right hemisphere works with semantic elements inside the proposition, the left hemisphere works with semantic elements outside propositions, either modifying them or putting them in some kind of mutual relations (causal, etc.). This understanding is fully supported by the reasoning experiments. The right hemisphere appears to control the quality of information within one proposition—e.g., the truthfulness (denotative accuracy) of premises (that is, units equivalent to a proposition), testing them against the realities of the world and/or personal knowledge/experience. The left hemisphere is responsible for the correctness of operational mechanisms (formal correctness of inferences), and also for semantically combining together into a single unit a set of connected propositions (two premises and a conclusion-leading question). This means that the right hemisphere works holistically only with respect to a proposition as a unit (linguistically speaking, a first-order predicate with its arguments), whereas the left hemisphere works holistically with a set of rules of structural organisation of several propositions (not going inside the proposition)

> and perceives the new cluster as a unity. So both hemispheres carry on 'holistic' functions, but their holicity is different in principle.

This conclusion has to be reconciled with more general ideas of the functioning of the hemispheres: the right hemisphere is supposed to operate cognitively with unified configurations, whereas the left one is supposed to process discrete items (Witelson 1987). Evidently the boundaries of the unified configurations, from the linguistic perspective, are restricted to entities which can be perceived as a) a recognized unit of any linguistic level, independent of how many elements of the lower level constitute it and of what grammatical rules were involved in formation of this unit (the highest unit so far was considered to be a proposition actualising a single event/situation) and b) a content familiar as a whole to the speaker. So the ability to work holistically in linguistic terms is the ability to recognize some entity as a unity corresponding to one meaning, without going into the detail of what were the rules which created this unity (hence the ability of the right hemisphere to understand idioms and metaphors). The left hemisphere is primarily processing the rules of formation of more complex units from more simple ones, so its work is connected with assembling the modules.

But the experiments with reasoning modes discovered that it is the right hemisphere which cannot perceive the syllogism holistically, whereas the left hemisphere can. How to explain this? Evidently, discourse theory provides the best explanatory framework. To be able to perceive such a complex structure as a syllogism consisting of a row of propositions as a unified entity (without being able to apply rules, since the right hemisphere cannot do that), there must be a prior familiarity with this script as a whole. But since this script does not exist (it is very problematic whether it can exist at all, considering the necessity of establishing an inter-propositional connection), this set of propositions does not constitute a single unit. On the other hand, the left hemisphere can build such a script because it can assemble modules as large as clauses, and so can perceive the resulting entity as a unit. But the meaning of this unit is only inferential, not denotative. So if the syllogism constitutes a script with a content (as was assumed in the discourse theory of reasoning) it is a very special 'syntactic script'—a system of formal rules with the function of those rules, which is never tied to a definite content but only to a definite set of rules. In my opinion, the latter understanding of the syllogism is much more plausible.

REFERENCES

Balonov, Lev Ja. & Vadim L. Deglin. 1976. *Hearing and language of dominant and non-dominant hemispheres*. Leningrad: Nauka. [in Russian]

Černigovskaja, Tatyana V. & Vadim L. Deglin. 1990. Syllogism solving under conditions of suppression of either the right or the left brain hemispheres. *Human Philosophy* 16:21–28. [in Russian].

Cole, Michael, John Gay, Joseph A. Glick & Donald W. Sharp. 1971. *The cultural context of learning and thinking*. New York: Basic Books.

Cole, Michael & Sylvia Scribner. 1974. *Culture and thought*. New York: John Wiley.

Dolinina Inga B. in press. Theoretical and empirical reasoning modes from neurological perspective. *Argumentation*.

Deglin Vadim L. 1995 (manuscript). Thought heterogeneity and brain hemispheres. 62 pages.

———, Lev Y. Balonov & Inga B. Dolinina. 1983. The language and brain functional asymmetry. In *Proceedings of the Tartu State University (Transactions on sign systems XVI)* 635:31–42. [in Russian].

——— & Marcel Kinsbourne. 1996. Divergent thinking styles of the hemispheres: how syllogisms are solved during transitory hemisphere suppression. *Brain and cognition* 31: 285–307.

Inhelder, Bärbel & Jean Piaget. 1958. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Ivanov, Vyačeslav V. 1978. *Even and odd: asymmetry of brain and of semiotic systems*. Moscow: Soviet Radio. [in Russian].

Jakobson, Roman. 1990. Brain and Language In *On language: R. Jakobson*, ed. by Linda R. Waugh & Monique Monville-Bursto, 499–513. Cambridge ma: Harvard University Press.

Johnson-Laird, Philip N. 1983. *Mental modes: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

——— & Ruth M. Byrne. 1991. *Deduction*. Hove: Lawrence Erlbaum.

Lévy-Bruhl, Lucien. 1923. *Primitive mentality*. New York: MacMillan.

Lévi-Strauss, Claude. 1966. *The savage mind*. Chicago: University of Chicago Press.

Lotman, Jurij M. 1978. The phenomenon of culture. *Proceedings of the Tartu State University (Semiotics of culture: Transactions of sign systems X)* 463:3–17. [in Russian]

Luria, Alexander R. 1976. *Cognitive development, its cultural and social foundations*. Cambridge ma: Harvard University Press.

Moore, Terence. 1986. Reasoning and inference in logic and in language. In *Reasoning and discourse processes*, ed. by Terry Myers, Keith Brown & Brendan McGonigle, 51-66. London: Academic Press.

Olson, David R., Nancy Torrance & Angela Hildyard (eds.). 1985. *Literacy, Language and learning: the nature and consequences of reading and writing.* Cambridge: Cambridge University Press.

Ong, Walter J. 1982. *Orality vs. literacy: The technologizing of the world*. New York: Methuen.

Piaget, Jean. 1971. The theory of stages in cognitive development. In *Measurement and Piaget*, ed. by Donald R. Green, Marguerite P. Ford & George B. Flamer, 1–11. New York: McGraw-Hill.

———. 1980/1964. *Six psychological studies*. Brighton: The Harvest Press.

———. 1985/ 1975. *The equilibration of cognitive structures*. Chicago: University of Chicago Press.

Scribner, Sylvia. 1977. Modes of thinking and ways of speaking: culture and logic reconsidered. In *Thinking: Readings in cognitive science*, ed. by Philip N. Johnson-Laird & Peter C. Wason, 483–500. Cambridge: Cambridge University Press.

——— & Michael Cole. 1981. The psychology of literacy. Cambridge ma: Harvard University Press.

Sharp, Donald W., Michael Cole & C. Lave. 1979. *Education and cognitive development: The evidence from experimental research*. [Monographs of the Society for Research in Child Development 44(2) (Serial #178)].

Wason, Peter C. & Philip N. Johnson-Laird. 1972. *Psychology of reasoning.* London: B.T. Batsford.

Wilson, Deirdre & Dan Sperber. 1986. Inference and implicature in utterance interpretation. In *Reasoning and discourse processes*, ed. by Terry Myers, Keith Brown & Brendan McGonigle, 241–64. London: Academic Press.

Witelson, Sandra. F. 1987. Neurological aspects of language in children. *Child development* 58:653–688.

~

DISCOURSE MARKERS—ACCESSING 'HEARER-OLD' INFORMATION: THE CASE OF RUSSIAN *ŽE*

Anna Feldman
Ohio State University/The Hebrew University of Jerusalem

INTRODUCTION. In this paper, I will examine the status of a discourse marker that has been claimed to be polysemous. I will demonstrate that all uses of the Russian discourse marker *že*, traditionally labeled as 'emphasis', 'justification', 'contrast' and even the diametrically opposing 'similarity' can, in fact, be reduced to a common denominator. I argue that despite the apparent versatility of *že*, this element realizes one discourse function; namely, it instructs the addressee to process the information in a certain discourse segment vis-à-vis 'hearer-old' information. The unified account I adopt here will utilize the framework provided by Relevance Theory.

1. PREVIOUS TREATMENTS OF ŽE: FORM AND FUNCTIONS. This section will sketch the previous treatments of *že* as a 'second position' clitic and as a 'polysemous' particle.

1.1. FORM: SECOND POSITION. The Russian *že* prototypically occurs after the first prosodic word, as in example (1), or constituent in the clause, as in example (2):

(1) *Čej* že *den' roždenja #že segodnja?*
Whose *že* birthday #že today
'So, whose birthday is it today?'

(2) *Počemu ty emu ne veriš? On* že *vrač.*
Why you him not believe? He *že* doctor.
'Why don't you believe him? After all, he is a doctor.'

Generative treatments have, thus, classified *že* as a 'second position clitic', a phenomenon that is commonly attested in almost all Slavic languages as well as in other languages. Indeed, *že* occurs in the second position in almost all instances of declaratives (example (3)), interrogatives (examples (4) and (5)) and imperatives (example (7)):

(3) *Maša* že *ne znala čto ja dolžen priexat'.*
Masha *že* not knew that I should come
'After all, Masha didn't know that I should come.'

(4) *Ty* že *byl v Kitae?*
You *že* was in China
'You have been to China, haven't you?'

(5) *Kakoje že plaťe #*že *ty kupila?*
What že dress #*že* you bought
'What dress did you buy?'
(6) *Otdai* že *mnje moi den'gi!*
Give-back *že* me my money
'Give me my money back!'

The analysis of *že* as a second position clitic apparently constitutes the correct descriptive characterization. *Že* follows either the first constituent of a clause (e.g., (2), (3), and (4)) or the first prosodic word (e.g., (1), (5) and (6)). However, such a proposal of an essentially syntactic-phonological nature, regarding the clitic distribution, does not shed light on what factors trigger the first prosodic word to be a host and what factors trigger the first constituent to be a host. Thus, for example, it is unclear why (1) and (5) are infelicitous when *že* occurs after the first syntactic constituent and not after the first prosodic word. The analysis of *že* as a clitic that occupies the second syntactic/prosodic position in a clause is insufficient in many respects. First of all, there are cases where *že* occurs later than the second position, as illustrated in the following examples:

(7) *My dolžny tam byť segodnja* že *večerom*
We must there be today *že* evening
'We have to be there tonight!'
(8) *U menja esť takaja* že *kniga*
I have such *že* book
'I have exactly the same book'.

In (7) and (8), *že* does not appear to occupy the second position, whichever way this position is defined. Thus, it is evident from these examples that the distribution of *že* should be defined in a more refined way, otherwise it does not predict all the occurrences of *že*.

Moreover, there are constructions that do not allow *že*. One such example is demonstrated in (9):

(9) *#Ty že podpisyvaešsja na 'Pravdu' ili 'Jersualem Post'?*
You *že* subscribe to 'Pravda' or 'Jerusalem Post'
'Do you subscribe to 'Pravda' or to 'the Jerusalem Post'?

(9) illustrates the impossibility of using *že* in alternative questions. However, if the distribution of *že* is restricted only to the first syntactic constituent or first prosodic word in clauses, the infelicity of this example remains unexplained.

The aforementioned examples suggest that the characterizing *že* as a second position clitic over-generates or under-generates the distributional properties of *že*. Thus, constructions, such as (7) and (8), where *že* apparently does not occupy the second position, are under-generated; whereas cases, such as (5) or (9) are over-generated. Their distributional properties are not predicted by the 'second position' rule. Moreover, adopting the analysis of *že* as a second position clitic, raises the issue of what constitutes the first position. The 'second position' analysis in purely syntactic terms is not sufficient, as has been explicitly exemplified; and in fact, examination of a wider range of data suggests that this phenomenon may have to be linked to information structure in general and the concept of information unit in particular. It should be mentioned that analyses that have been proposed in an attempt to solve the problem of 'the second position', had to make a variety of ad hoc assumptions (e.g., Barbosa's 1996 and Hock's 1996 'skipping', Taylor's 1996 'prosodic inversion' etc.). The syntactic-prosodic analysis of *že* neither accounts for the distribution of this element nor predicts all its occurrences. In section 4, I will demonstrate how the occurrence of *že* in positions other than second in the clause (as exemplified in (7) and (8)) can be incorporated into a 'second position' analysis.

1.2. FUNCTION: THE TRADITIONAL FOUR KINDS OF ŽE. As has been mentioned, the discourse marker *že* is highly frequent in Russian. Traditional grammarians (e.g. Rozental and Telenkova 1985 and others) treat *že* as a 'polysemous particle' and characterize the range of meanings in terms such as 'emphasis', 'contrast' and even the completely opposite 'similarity', as well as providing examples of the type that I refer to as 'justification'. The range of possible meanings that were enumerated by traditional grammarians is demonstrated in the following examples:

A. CONTRAST. Examples such as (10) and (11) are treated by traditional grammarians as contrastive. Accordingly, it is *že* that contributes a so-called contrastive reading to these utterances.

(10) *On ostaetsja, ona* že *uezžaet*
He stays, she *že* leaves
'He is staying, [but] she, after all, is leaving
(11) *Maša umnaja, Daša* že *net*
Masha smart, Dasha *že* not
'Masha is smart, [but] Dasha, after all, is not'

B. JUSTIFICATION. Another meaning of *že* that I refer to as 'justification' is compatible with the English *after all* or Hebrew *harey.*

(12) *Čto ty stoiš? Sadis' v mašinu. Ona* že *naša.*
What you stand? Get into car. It *že* ours.
'What are you waiting for? Get into the car. It's ours.'
(13) *Potoropis'! On* že *tebja ždet!*
Hurry up! He *že* you wait
'Hurry up! He's waiting for you, after all!

C. EMPHATIC READINGS. There are certain uses of *že* that traditional grammarians relate to as 'emphatic'. What this term designates is unclear. In the following examples, *že* is treated as having a so-called 'emphatic' reading. In these utterances, *že* is analogous to the Hebrew *kvar* ('already') or *nu* whose function in certain contexts can be defined as 'urging' or 'driving forward'. The following examples, according to the traditional approach, display an emphatic function:

(14) *Kogda* že *ty datbereš v komnate?*
When *že* you will clean in room
'When will you clean your room?'
(15) *Skaži* že *mne čto-nibud'!*
Tell *že* me something
'Say something already!'
(16) *My dolžny byť tam segodnja* že *večerom!*
We must be there today *že* evening
'We have to be there tonight!'
(17) *Čto* že *mne delať?*
What *že* me to do
'What shall I do then?'
(18) *Kto* že *znal čto pogoda yizmenitsja?*
Who *že* knew that weather change
'Who could have known that the weather would change?

D. SIMILARITY. Sometimes, *že* is described as conveying the 'similarity' meaning:

(19) *U menja esť takaja* že *kniga*
I have such *že* book
'I have exactly the same book.'
(20) *V tu* že *minutu razdalsja zvonok*
At that *že* moment rang bell
'At that very moment the bell rang'.

The inadequacy of a polysemous account is evident from a comparison of instances occurring with and without *že*. Thus, (21)a and b exemplify the 'contrastive' reading:

(21) a. *On ostaetsja, ona* že *uezžaet*
He stays, she *že* leaves
'He is staying, [but] she, after all, is leaving.
b. *On ostaetsja, ona uezžaet*
He stays, she leaves
'He is staying, [but] she is leaving.

Clearly, the contrastive meaning is not contributed by *že* itself. Even without the particle *že*, these examples are contrastive, because the contrast is present in the semantics of the utterances, as indicated in (21)b. Therefore, the 'contrastiveness' cannot be characterized as an inherent property of this element.

A similar claim can be made about the 'justification' and 'similarity' readings, as is demonstrated in (22)a and b, and (23)a and b, respectively:

(22) a. *Čto ty stoiš? Sadis' v mašinu. Ona* že *naša.*
What you stand?Get into car. It *že* our car.
'What are you waiting for? Get into the car. It's ours.'
b. *Čto ty stoiš? Sadis' v mašinu. Ona naša.*
What you stand? Get into car. It ours.
'What are you waiting for? Get into the car. It's ours.'

The justification meaning is preserved when *že* is omitted, as in (23b).

(23) a. *U menja est' takaja* že *kniga*
I have such *že* book
'I have exactly the same book .'
b. *U menja est' takaja kniga*
I have such book
'I have such a book .'

(23)b can be interpreted as displaying the 'similarity' meaning. Thus, like 'contrast', 'justification' and 'similarity' are not inherent properties of *že* either.

As for the term 'emphatic', it is problematic too, since it covers a variety of uses with no apparent common denominator as in (17) and (18) that cannot be clearly characterized in terms of emphasis. No parameters for defining the emphatic function are provided and the term is left loose and open-ended. As in the previous cases, the elimination of *že* will result in no change of propositional content.

(17) *Čto* že *mnje delat'?*
What *že* me to do
'What shall I do then?'
(18) *Kto že znal čto pogoda izmenitsja?*
Who *že* knew that weather change
'Who could have known that the weather would change?

In sum, the treatment of *že* as a polysemous element that has a variety of totally unrelated uses has been shown to be inadequate in accounting for the range of data. The list of uses of *že* is not exhaustive. As has been shown, some examples cannot be classified under any of the given labels (e.g. the so-called cases of emphasis). Likewise, the fact that totally arbitrary and completely opposed functions, such as 'similarity' and 'contrast', have been associated with *že* suggests that a more insightful generalization concerning its meaning is required.

As previously noted, in all of the given examples, the different readings are present in the propositional content of the utterance, with or without *že*. Thus, *že* does not contribute to the propositional content of the utterances. This leads to the possibility that an analysis of *že* should examine its contribution to the organization of the information structure, rather than to the semantics or propositional content of the utterance. Along these lines, a unified analysis of *že*, accounting for the variety of uses, will be proposed within the framework of Relevance Theory.

2. MULTI-FUNCTIONALITY AND COMMON DENOMINATOR. In this section, the Russian discourse marker *že* is examined within the conceptual/procedural dichotomy proposed by Sperber and Wilson (1993) in the framework of Relevance Theory. The specific procedural function of *že* is argued to be both relevance-motivated and structurally determined. The relevance-theoretic framework allows for a unitary account for the seemingly disparate uses of *že*.

2.1. *ŽE* IN RELEVANCE-THEORETIC TERMS. As will be shown in this section, despite the apparent versatility of *že*, this element realizes a unified discourse function; namely, it instructs the addressee to process the information contained in the utterance segment vis-à-vis activated or re-activated information that occurs in the given linguistic or extra-linguistic context.

Before presenting the Relevance-theoretic account of *že*, the necessary details of the general theoretical framework will be outlined.

2.1.1. RELEVANCE THEORY AND CONCEPTUAL/PROCEDURAL DICHOTOMY. Relevance Theory, proposed by Sperber and Wilson (1986), is based on the assumption that human understanding is guided by relevance considerations, which control the inferential processes involved in interpretation. Sperber and Wilson's (1986) fundamental idea is that hearers are bound to interpret utterances as maximally relevant where relevance is regarded as a trade-off between contextual effects (i.e. informativity) and processing costs. The role of the context is to supply assumptions that are brought to bear in the assessment of relevance. This means that expressions that constrain the hearer's choice of context must be analyzed as linguistic (or semantic) constraints on relevance. The speaker's task is to create an expectation that the utterance will achieve enough contextual effects to be worth processing for the hearer and cause him no unnecessary processing effort (optimal relevance). The hearer's task is to assess which effects the speaker could have intended to achieve, in other words, to constrain a context which will make the utterance worth processing.

Sperber and Wilson (1986) proposed a distinction between conceptual and procedural elements. Conceptual meaning encodes information that contributes to the propositional content and prototypically truth-conditionally relevant, whereas procedural meaning encodes processing instructions, signaling to the addressee how to process the utterance and usually does not contribute to the propositional content of utterances.

Discourse markers contribute to relevance by operating as signals that tell the hearer how an utterance is to be understood, thus, reducing the processing efforts. Accordingly, the discourse marker *že* is a procedural element, signaling to the addressee that the speaker takes certain information in the utterance to be hearer-old (in the sense of Prince (1981), Birner & Ward (1998)), and instructing the addressee to process the information contained in the utterance segment vis-à-vis activated or re-activated information. Prince (1981) offers a pair of crosscutting dichotomies which classify information as, on the one hand, either 'discourse-old' or 'discourse-new'; and on the other hand, either 'hearer-old' or 'hearer-new'. This distinction is essential in accounting for various distributional constraints, such as the so-called definiteness effect in existential sentences. Discourse-old information is information evoked in the current discourse; hearer-old information is the information that the speaker believes to be present within the hearer's knowledge store. What is new to the discourse is not necessarily new to the hearer. However, what is 'discourse-old' should be necessarily 'hearer-old'. Constructions may vary with respect to whether they are sensitive to discourse-familiarity or hearer-familiarity. The element *že*, thus, indicates to the addressee that the information contained in the current utterance or activated by it is expected to be 'hearer-old', thereby restricting the context used in processing.

2.1.2. THE UNIFIED ACCOUNT IN RELEVANCE THEORY. In this section, the variety of instances of the discourse marker *že* is analyzed. The utterances in (12) with and without *že* as in the following have different contextual effects:

(24) a. *Čto ty stoiš? Sadis' v mašinu. Ona* že *naša.*
What you stand? Get into car. It *že* ours.
'What are you waiting for? Get into the car. It's ours.'
b. *Čto ty stoiš? Sadis' v mašinu. Ona naša.*
What you stand? Get into car. It ours.
'What are you waiting for? Get into the car. It's ours.'

In (24)b, the information that the car belongs to the speaker is not taken to be 'hearer-old'. In this case, the speaker may merely be informing the addressee that this car belongs to him/her. A possible situation where (12)b can be uttered is a context where the speaker who bought a new Porsche without letting his wife know points to the car uttering *Čto ty stoiš? Sadis' v mašinu. Ona naša.* In this context the *že*-utterance will be infelicitous, since the fact that the new Porsche belongs to the family is completely new to the hearer (the wife). In (24)a with *že*, on the other hand, the fact that the car is 'ours' is assumed by the speaker to be in the hearer's store of knowledge. Such an utterance will be natural in the following context: looking for a car at the parking lot and eventually having found it , the husband says to his wife who is still not recognizing her property: *Čto ty stoiš? Sadis' v mašinu. Ona že naša. Že* is only felicitous in a context where the speaker assumes that this information is already known or should be known to the hearer. In this case the information may be retrieved from long-term memory, from the situational context or the linguistic context. Consider (25)a and b:

(25) a. *Potoropis'! Oni* že *tebja ždut!*
Hurry up! They *že* you wait
'Hurry up! They are waiting for you, after all!
b. *Potoropis'! Oni tebja ždut!*
Hurry up! They you wait
'Hurry up! They are waiting for you!

(25)a can occur without prior linguistic context. For example, in a theatre, the audience is waiting for one of the actors, who is still busy with his make-up. The actor is likely to be told: 'The audience *že* is waiting for you!' This context explicitly shows that the information within the utterance is given to the addressee or at the very least is assumed by the speaker to be given. The same linguistic sequence without *že* will neutrally convey the information that they are waiting/ or the audience is waiting that might or might not be new to the hearer; however, when *že* occurs in this utterance, the information must be 'hearer-old'. All instances of *že* can be analyzed along these lines.

(10) is felicitous in a context where the hearer is supposed to know or to be able to infer that the person referred to is leaving. By using *že*, the speaker intends to signal that the addressee is expected to have the relevant information in his store of knowledge. A possible context for (10) might be a situation like the following:

A: Are they both leaving? B: He is staying, She *že* is leaving.

Such a context explicitly shows that the information that she is leaving is not new for the hearer. However, if the speaker chooses to inform the hearer that 'he is staying, she is leaving' without the previous assumption that the information is 'hearer-old', the utterance with *že* will be infelicitous.

(14)—(18) which were treated as indicating emphasis can be dealt with in a similar fashion. In all of them, *že* is used to instruct the addressee to process the information contained in the utterance segment vis-à-vis activated or re-activated information that occurs in the given linguistic or extra-linguistic context. (14) is felicitous in a context like the following:

> A child was asked to clean his room. It is likely to happen that when his mother walks into the room and sees that the room is a mess, she will utter 'When *že* will you clean your room?'

This context clearly demonstrates that *že* can be used only when the information modified by this element is 'hearer-old'. To the kid the information that he should have cleaned the room is not new.

From the examples discussed so far, it is evident that the Russian entity *že* functions as a unified signal to the addressee to process the information contained in the utterance segment in question vis-à-vis activated or re-activated information that occurs in the linguistic or extra-linguistic context.

I will now address the 'second position' issue.

2.1.3. FORM – 2ND POSITION RELATIVE TO SCOPE. As has been mentioned earlier in the paper, in previous literature, *že* was treated as a second position particle. In (7), (8), and (1), the position occupied by *že* is neither the second prosodic word, nor the second syntactic position in the clause. Note that in (7), the Russian expression for 'tonight' consists of two words 'today evening'. A possible context for (7) might be the following:

a. *Masha*: When are we going?
b. *Ivan*: We are going tonight. (in Russian 'today evening')
c. *Masha*: Let's wait for Gosha and go tomorrow.
d. *Ivan*: We have to be there tonight že (in Russian 'today *že* evening')

In (d), *že* modifies a domain smaller than the clause. Using *že*, the speaker explicitly indicates that the information that they are going tonight may be retrieved from the context (in our case from the linguistic context, specified in **b**). The relevant information that should be reactivated in this case is not the whole clause 'we are going tonight', but only the constituent 'today evening'. In order to express the retrievability from context of 'we are going tonight', the speaker would use *že* following 'we' in **b**: 'We *že* are going tonight'.

To illustrate further that *že* can have a domain smaller than the clause, consider context (8)' for the utterance in (8).

(8)' a. *A*: I don't believe it, I have such a book!
b. *B*: I am not sure that this is the same book.
c. *A*: No, I bought such *že* a book. (i.e., I bought exactly the same book)

In this example again, the information that should be retrieved from the hearer's store of knowledge is not that, contained in the whole utterance in (8)'a, but the constituent 'such *že* a book'. The event of 'A's buying a book', expressed in the predication in (8)'c, might be unknown to the hearer, but the fact that A has exactly the same book is evident from the linguistic context.

It should be mentioned that (8) can occur without prior linguistic context. In a context, where the speaker sees that a book, lying on the table, is exactly the same as his, he is likely to utter (8). It is thus evident that the information, which is taken to be re-activated, need not be linguistically specified, but could be provided situationally.

(1) is another example where *že* seems to modify not the whole clause but only a single constituent—in this case again an NP. A possible context for (1) might be the following: During a birthday party, a person suddenly asks one of her friends:

a. *A*: Whose birthday is it today? Mary's?
b. *B*: No.
c. *A*: Peter's?
d. *B*: No.
e. *A*: Whose *že* birthday is it today? (The Russian variant of this utterance is given in (1))

The difference between the question without *že*, in **a**, and its counterpart with *že*, in **e**, is again in the contextual effects of these utterances. In the question without *že*, the speaker has a presupposition that this is somebody's birthday, because it is a birthday party. Such a presupposition is based on the knowledge of the world. In the case with *že*, in addition to this presupposition, the speaker makes the assumption that there exists a restricted set of people whose birthday it might be and that this set in question is available to the hearer. The possible felicitous responses to this question are restricted. They have to include members of the relevant set. The question in (a), without *že*, does not indicate that the hearer must choose one of the people from a given group. Thus, in addition to the regular effect of the wh-question, where the speaker assumes some set of information options to be given, *že* restricts a given set to a given proper set.

It appears that expressions, such as those in (1), (7) and (8), are not exceptions. They can be subsumed under the 'second position' analysis. As can be easily observed, the discourse marker *že* occupies the second position in a relevant sub-segment.

Such an effect is characteristic of other constructions, such as wh-clefts and conditionals.

According to Prince (1978), a wh-cleft occurs coherently in a discourse only if the material inside the (subject) wh-clause represents material which a cooperative speaker can assume to be appropriate in the hearer's consciousness at the time of hearing the utterance. In relevance-theoretic terms, adopted for our analysis, this material is already hearer-old information. When *že* occurs in such constructions, it further restricts the context, i.e., this element signals that the information modified by *že* is properly included in the given relevant set of information alternatives which is hearer-old. Consider the following example:

(26) a. *To čto ja xoču skazat'—sovsem ne to čto vy dumaete.*
what I want to say—completely not what you think
'What I want to say is completely different from what you think.'

(26) b. *To čto že ja xoču skazať—sovsem ne to čto vy dumaete.*
what *že* I want to say—completely not what you think
'What I want to say is completely different from what you think'.

The difference between (26)a and b is in their contextual effects. (26)a indicates that the hearer knows the speaker is going to say something. In addition to this effect, *že* in (26)b signals that the hearer not only knows that the speaker is going to say something, but he knows the speaker could have said a number of things but he chose only one alternative from the relevant set. (26)b might occur in a context where the speaker utters: 'You, of course, think that I am going to give you a lecture on morality, or maybe you think that I will read you a passage from the Bible, but what *že* I want to say is completely different from what you think'. The speaker who uses a wh-cleft with *že* instructs the hearer to infer that the choice of the particular set in question is available, even though the addressee does not know what exactly the speaker intends to say.

Že functions in a similar fashion in conditionals. Examine (27)a and b:

(27) a. *Yesli ja vypju sejčas kofe, ja ne budu spať noč'ju.*
If I drink now coffee, I will not sleep at night
'If I drink coffee now, I will not sleep at night.'
b. *Yesli* že *ja vypju sejčas kofe, ja ne budu spať noč'ju.*
If *že* I drink now coffee, I will not sleep at night
'If I drink coffee now, I will not sleep at night.'

The utterances given in (27)a and b differ from each other in that they might occur in distinct contexts. The condition in (27)a might or might not be new information to the addressee. However, (27)b indicates that the hearer is supposed to infer that there exists a set of conditions, and that the condition in the utterance is one of the relevant alternatives. The interaction of the function of *že* with the general function of conditional sentences will result in an additional contextual effect. Thus, (22)b might appear in a context like the following: 'If I drink hot milk now, I will fall asleep and will not be able to write my paper, but if *že* I drink coffee now, I will stay awake for the entire night'.

To sum up, as has been shown, *že* instructs the addressee to process the information contained in the relevant utterance segment vis-à-vis activated or reactivated information that occurs in the linguistic or extra-linguistic context. As has been illustrated, *že* is only felicitous in a context where the speaker assumes that relevant information is already known or could be available to the addressee from the previous discourse or from the non-linguistic context (either situational or based on the knowledge of the world).

I have drawn upon the traditional analysis of *že* as a second position particle, extending it to account for those cases in which *že* does not seem to occupy the second position. I have shown that *že* can take scope over constituents smaller than the clause, and in these cases, it occurs in the second position within this constituent. Thus, these apparently exceptional cases can be subsumed under the second position analysis.

In adopting the second position analysis for *že*, the question that emerges concerns the nature of the first position. It has been shown that *že* can refer to the whole utterance as well as to a relevant sub-segment.

3. FURTHER PREDICTIONS. The analysis proposed in this paper makes the correct predictions with respect to such constructions as yes-no questions and alternative questions. Thus, for instance, if *že* occurs in a polar interrogative, it will interact with the general function of this kind of questions and the result will be a rhetorical question or a request for confirmation.

(28) a. *Ty* že *podpisyvaešsja na 'Jerusalem Post'?*
You *že* subscribe to 'Jerusalem Post
'You subscribe to 'the Jerusalem Post, don't you/ right?'
b. *Ty podpisyvaešsja na 'Jerusalem Post'?*
You subscribe to 'Jerusalem Post
'Do you subscribe to the Jerusalem Post?'

Že indicates that the speaker assumes that the information modified by *že*, i.e. that X subscribes to the Jerusalem Post is known. The utterance in (28)a can appear in a context like the following:

Masha to Ivan: I have no idea who Putin is. Is he still working for the KGB?
Ivan to Masha: You *že* subscribe to 'the Jerusalem Post', don't you? Why don't you ever read it?

The utterance in the context indicates that the information contained in the question is already taken by the speaker to be in the hearer's store of knowledge. The basic function identified for *že* can account for such cases as well. *Že* instructs the addressee to process the information contained in the utterance segment vis-à-vis activated or re-activated information that occurs in the given linguistic or non-linguistic context (in this case, knowledge of the world). The interaction of the function of *že* with the general function of yes-no questions will result in rhetorical questions or in request for confirmation, as demonstrated in (28)a. (28)b, which does not contain *že* will occur in another context. For example, 'I would like to see the latest article on Putin. Do you subscribe to the Jerusalem Post?' In this context, the speaker does not convey an assumption that he already knows the hearer subscribes to the Jerusalem Post.

Another prediction that can be made on the basis of the analysis proposed here is that *že* will not occur in alternative questions, as illustrated in (29):

(29) # *Ty* že *podpisyvaešsja na 'Pravdu' ili 'Jerusalem Post'?*
You *že* subscribe to 'Pravda' or 'Jerusalem Post'
'Do you subscribe to 'Pravda' or 'the Jerusalem Post'?

Under the proposed analysis this fact follows naturally. The impossibility of using *že* in example (29) is predictable. *Že* will be infelicitous in a context where the speaker, uttering an alternative question, has no prior expectation as to which alternative will be chosen. The freedom to choose between two or more alternatives blocks the function of *že* as a marker of known/inferable information which should be processed vis-à-vis activated or re-activated information. Alternative questions contain mutually exclusive information and therefore, will never co-occur with *že* that takes scope over information known but not salient at the given point, as illustrated in (29). However, this question becomes felicitous, if we take the sequence 'Pravda or the Jerusalem Post' in a context, where the addressee is not supposed to choose between two or more things; for example, if that speaker merely wants to make sure that the hearer indeed subscribes to one of these newspapers, no matter which one. In this case, then 'Pravda or the Jerusalem Post' is taken as one item of information and the question is then not treated as alternative. The information that the hearer subscribes to one of these newspapers without necessarily specifying which one exactly is taken to be 'hearer-old', and thus, this question functions as request for confirmation.

4. CONCLUSION. I have proposed a unified account of *že* as a discourse marker that instructs the hearer to process the information contained in the utterance or in a relevant sub-segment vis-à-vis activated or re-activated information that occurs in the given linguistic or extra-linguistic context. The study of the Russian discourse marker *že* indicates that adopting Relevance Theory considerations affecting discourse structure with a distinction between conceptual and procedural uses of lexical items makes a coherent and unified characterization of this discourse marker possible.

The previous analyses in syntactic-prosodic terms have been shown to be unable to explain the non-occurrence of *že* in alternative questions as in (29). They also were incapable of accounting for the occurrence of *že* in positions that in terms of linearity could not be described as second, as in (7) or (8). It has been shown that the scope of *že* is insightfully accounted for by its position in utterances, defined in terms of linearity and information units.

The fact that *že* is given a unitary account within a theoretical framework that reconciles Relevance factors and structural consideration, raises a question for further investigation as to its applicability in cross-linguistic analyses of discourse markers in general. The question whether such an analysis indeed applies insightfully to other discourse markers and provides a coherent definition of discourse markers awaits further research. Since codification of information as 'hearer-old' is evident cross linguistically, and in most cases is codified lexically, comparative studies could thus increase our understanding of the division of labor between the various discourse markers in different languages.

---

[1] The term 'justification' does not appear in the traditional textbooks. The traditional descriptions provide examples of the type that I refer to as 'justification'. Most dictionaries consider English *after all* or Hebrew *harey* as the best translations of *že*.

[2] I use the terminology proposed by Halpern (1992, 1996), Inkelas (1989), Marantz (1988), Progovac (1996) and others, who differentiate two kinds of so-called 'second position': after the first prosodic word or after the first syntactic constituent. This terminology is used for Serbo-Croatian, Polish, Greek and other languages.

[3] It is difficult to translate utterances with *že* into English. English cannot codify certain uses of *že* in an exact way.

[4] In the following sections cases where *že* does not appear second, whichever way this position is defined, will be included in the 'second position' analysis.

[5] Note that in such examples the use of *že* can be regarded as attitudinal, since it codifies some attitude conveyed by the speaker.

[6] Principle of Relevance: Every act of ostensive communication communicates a presumption of its own optimal relevance (Sperber & Wilson 1986).

[7] In the light of Sperber and Wilson (1993), studies utilizing the conceptual/procedural distinction have indicated the need for some modification in this strict division and have claimed that while it is feasible to maintain the linguistic/non-linguistic division, this does not correspond to the separation between truth-conditional and non-truth-conditional aspects of propositions.

[8] It should be mentioned that the codification of information as 'hearer-old' is evident cross-linguistically. Thus, for example, the Hebrew partial correlates of *že* are *nu*, *mamash*, *kvar*, *harey*, etc., whereas the English partial correlates of the Russian discourse marker are *after all*, *already*. The following examples indicate the relevant 'hearer-old' codification:

a. Why don't you believe him? He *že/harey/after all* is a doctor.
b. Come *že*! / Come *kvar!/ Nu*, come!/ Come on *already*! *She* is waiting for you!
c. I have such *že* a book./ I have *exactly* the same book/ I have *mamash* the same book.

[9] The effect of the interaction between the position of *že* and various intonation clues is left for further research.

REFERENCES

Barbosa, Pilar. 1995. Clitic placement in European Portuguese and the position of subjects. In Halpern & Zwicky. Stanford: CSLI Publications.

Birner, B. & G. Ward. 1998. *Information status and noncanonical word order in English*. Amsterdam: John Benjamins

Franks, Steven 1998. Clitics in Slavic. Position paper presented at Comparative Slavic Morphosyntax. Spencer, Indiana, 5–7 June 1998.

Halpern, Aaron L. & Arnold M. Zwicky. 1996. *Approaching second: Second position clitics and related phenomena.* Stanford: CSLI Publications.

Hock, Hans H. 1996. Who is on first? Toward a prosodic account of P2 clitics. In Halpern & Zwicky.

Inkelas, Sharon. 1989. Prosodic constituency in the lexicon. Doctoral dissertation: Stanford University.

Jucker, H. Andreas & Yael Ziv. 1998. Discourse markers: Descriptions and theory, Amsterdam: John Benjamins.

MARANTZ, ALEC. 1988. Clitics, morphological merger, and the mapping to phonological structure. In *Theoretical morphology*, ed. by M.Hammond & M.Noonan. New York: Academic Press.

PRINCE, ELLEN F. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language* 54:883–906.

———. 1981. Toward a taxonomy of given-new information. In *Radical pragmatics*. New York: Academic Press.

PROGOVAC, LJILJANA. 1996. Clitics in Serbian/Croatian: Comp as the second position. In Halpern & Zwicky.

ROSENTAL D.A. & M.A. TELENKOVA. 1985. *Slovar' trudnostei russkogo jazyka*. Moskow: Russkii Jazik.

SPERBER, DAN & DEIRDRE WILSON. 1986. *Relevance: Communication and cognition*, Oxford: Blackwell.

——— & ———.1993. Linguistic form and relevance. *Lingua* 90:1–25.

TAYLOR, ANN. 1996. A prosodic account of clitic position in ancient Greek. In Halpern & Zwicky.

* * *

ANCIENT THEORIZING ON NONVERBAL COMMUNICATION

Thorsten Fögen
University of Exeter (Great Britain)

introduction. Nonverbal communication has received attention in many disciplines. Apart from communication studies in the narrow sense, it forms one of the areas of interest in psycholinguistics, semiotics, language acquisition research, foreign language studies and rhetoric as well as in ethology, sociology and psychology[1]. It is striking that the majority of studies from these areas neglect the historical dimensions of nonverbal communication. There is of course no denying that reference is made to theoretical works of modern times such as John Bulwer's *Chirologia* and *Chironomia* (1644), Johann Jakob Engel's *Ideen zu einer Mimik* (1785) and Charles Darwin's *The expression of the emotions in man and animals* (1872)[2], but the fact that nonverbal communication was analysed in Greek and Roman antiquity from a more theoretical perspective has only been emphasized in the past few years, mainly with reference to the connection between rhetoric and gesture[3]. This is even more remarkable since there exists a number of archaeological investigations into the depiction of nonverbal communication on ancient vases and coins as well as in the form of sculptures[4]. It is the aim of my contribution to give a brief overview of the most important aspects of theoretical reflections on nonverbal communication in ancient writings. In doing so, I will attempt to highlight some vital perspectives for future research which I am currently pursuing in a more comprehensive project.

1. ELEMENTS OF NONVERBAL COMMUNICATION AS HUMAN UNIVERSAL LANGUAGE.

Nonverbal communication is referred to in ancient sources in a variety of contexts. Very often the authors in question write about situations in which people with different mother tongues come into contact and find themselves in need of a mode of communication. The ancient sources mention various circumstances: Communication functions unproblematically if one speaks the language of the other or if there are competent interpreters who can assist people in making themselves understood. In the Roman world, it is predominantly people from the educated upper-class who have a reliable knowledge of Greek, mainly because of their bilingual education in Latin and Greek[5]. Among the Greeks, however, other cultures and languages were pushed into the background to the extent that other peoples (and for a long time even the Romans) were regarded as 'barbarians' and their languages, accordingly, as barbarous idioms. Viewed like this, it is not surprising that the low degree of open-mindedness and receptivity for foreign languages among Greeks and Romans should often have resulted in communication problems with other peoples[6].

That linguistic diversity and the lack of ability to communicate by means of spoken language may lead to alienation among people and that it can prevent more intensive contact with others right from the beginning, is a point mentioned by Pliny the Elder (*Naturalis historia* 7.7) and, later, by Augustine. According to the latter, the feeling of being different from a person with another language might lead one to prefer the company of one's dog to that of a foreigner (*De civitate Dei* 19.7 = PL 41.633–34 Migne). Being forced into a position of non-communication on the grounds of lacking competence in foreign languages appears to have been a matter of embarrassment to the ancients. Thus, silence is best prevented by avoiding contact with foreigners[7]. In his *Tusculans*, Cicero underlines—interestingly starting from the Romans' nonexistent knowledge of Greek and the Greeks' lack of familiarity with Latin—that not to understand a foreign language is tantamount to being deaf (*Tusc.* 5.116).

In Roman antiquity, the Latin language was valued as an integral part of the unity of the Roman empire. Connected with this was not only common ground on moral, political and religious levels but also the idea of a linguistically distinct society. This is pointed out by Cicero in a passage of his work *De officiis* (1.53), where he attributes to the mother tongue a high integrative value for its speakers. In addition to their political and national affiliation, the common language plays a vital role for the development of their identity; it generates a feeling of solidarity, a sense of a common bond among its speakers[8]. At the same time, we can deduce from passages such as Cicero's that in his time there was already a well-developed awareness concerning linguistic diversity in the ancient world, even if on the whole the Romans did not pay much attention to foreign languages other than Greek.

In this context, it is remarkable that ancient writers often point out that forms of nonverbal communication such as gesture and facial expression not only serve native speakers as visual support for the spoken word, but also that because of their vividness and concreteness they can function as a kind of universal language. According to some authors, this is due to the fact that all human emotions have by nature a corresponding expression in face, voice and gesture—a hypothesis which is much disputed in modern research[9]. As Cicero notes:

> *Atque in eis omnibus, quae sunt actionis, inest quaedam vis a natura data; qua re etiam hac imperiti, hac vulgus, hac denique barbari maxime commoventur: verba enim neminem movent nisi eum, qui eiusdem linguae societate coniunctus est, sententiaeque saepe acutae non acutorum hominum sensus praetervolant: actio, quae prae se motum animi fert, omnis movet; isdem enim omnium animi motibus concitantur et eos isdem notis et in aliis agnoscunt et in se ipsi indicant.* (Cicero, *De oratore* 3.223)

> 'And all the factors of delivery contain a certain force bestowed by nature; which moreover is the reason why it is delivery that has most effect on the ignorant and the mob and lastly on barbarians; for words influence nobody but the person allied to the speaker by sharing the same language, and clever ideas frequently outfly the understanding of people who are

> not clever, whereas delivery, which gives the emotion of the mind expression, influences everybody, for the same emotions are felt by all people and they both recognize them in others and manifest them in themselves by the same symbols.'

In another passage Cicero compares elements of human nonverbal expression to a string instrument in terms of function and effect (*De oratore* 3.216). For a linguistically ill-prepared listener or a foreigner, gesture, facial expression and voice illustrate everything which the spoken word cannot express in a similarly clear way. These considerations lead Cicero to insist on the following requirement in particular for the professional orator: Even if an orator's talent for verbal expression is highly developed, he should not refrain from making skilful use of accompanying nonverbal elements in order to convince the whole range of his multi-faceted audience.

In a rhetorical context one has also to mention the treatment of nonverbal communication in Quintilian's didactic treatise *Institutio oratoria*, but, as in the case of Cicero, they also transcend the boundaries of instruction on how to use facial expression, gesture and voice. It is emphasized that the importance of these elements for rhetoric can be recognized from their effects in other areas. Quintilian argues that much information can be transmitted without words if one lets the body speak—it is mainly the hands and the face which serve as instruments of communication by mutes. Similarly, animals signal emotions such as wrath and joy by body language. A comparable practice is the expression of a state of mind by dancing (*Inst. orat.* 11.3.65ff.; cf. also 11.3.87 with reference to the hands as the 'universal language' of man), which is, of course, a topic of a treatise on the defence of this art (*De saltatione*) written more than half a century later by the Greek author Lucian[10]. According to Lucian, the pantomimic dance portrays different characters and emotions (*De salt.* 35ff., cf. also 67), especially figures from mythology (cf. the long list of themes in *De salt.* 37–61). Like the orator the dancer must attempt to achieve a maximum of clarity (σαφήνεια) in his performance in order to make it comprehensible to his audience (*De salt.* 62). Dancing is said to unite pleasure (τὸ τερπνόν) and usefulness (τὸ χρήσιμον) in a unique way because it impresses the spectators not only by its harmonious movements but at the same time teaches them about the past (*De salt.* 71). A foreign guest of the Roman emperor Nero particularly admired an extremely talented pantomime dancer who lived at his court. While this foreigner did not understand the text of the song that accompanied the dance of the pantomime, it was the dancer's highly illustrative performance which permitted him to follow the entire plot without difficulties. When Nero asked the guest what he wished as a farewell gift, the foreigner opted for the dancer who by his nonverbal art would serve him as an interpreter in negotiations with other peoples (*De salt.* 64). Athenaeus, another Greek writer, reports on another dancer called Memphis who explained Pythagorean philosophy more clearly by his dancing than others by the use of words (1.20 C–D). It seems quite likely that such statements go back to or are at least inspired by the treatment of the dance in Plato's *Laws* (esp. 2.653 D9–654 A3, 2.672 D, 2.673 C11–D5, 7.816 A3–7; cf. also 7.814 D10–815 D5).

That speaking with one's hands is sometimes the only way to make oneself understood to foreigners was an experience which the poet Ovid had when he was banished to Tomis, nowadays the Romanian seaside resort Constanța (on this more extensively Fögen 2000:37ff.). According to his own account, he communicated with the inhabitants of Tomis by gestures:

exercent illi sociae commercia linguae:
per gestum res est significanda mihi.
barbarus hic ego sum, qui non intellegor ulli. (*Tristia* 5.10.35–37)

'They hold intercourse in the tongue they share;
I must make myself understood by gestures.
Here it is I that am a barbarian, understood by nobody.'

Admittedly these verses as well as the entire description of the Black Sea in Ovid's two collections of poems *Tristia* and *Epistulae ex Ponto* have to be judged as an exaggeration based upon the literary topos of the *locus horribilis*. Distrust of Ovid's portrayal of Tomis is even more apposite as on other occasions he refers to his study of the Getic and Sarmatian language and even to poetic works in Getic from his own pen. Nonetheless, the passage cited contains a constant trait of the Roman view of the mother tongue and foreign languages. Here as in other sources, we find expressed an attitude which is still very widespread in the first century A.D.: in regarding Rome, the capital of the empire, as the centre of the world, many Romans looked down on everything non-Roman; this includes languages with the exception of Greek.

Gesture and facial expression are similarly classified as universal language in the context of first-language acquisition. In his *Confessions*, Augustine conveys some highly informative insights concerning his own linguistic socialization (*Conf.* 1.6.7–1.19.30)[11]: As a baby (*in-fans*: 'speechless'), he tells us, he explained his wishes and desires to his environment by weeping, smiling, thrashing and screaming—by signs, as he adds himself, which corresponded to his wishes. However, such signs were not sufficiently unambiguous and sometimes caused misunderstandings among the adults around him; when they disobeyed him, little Augustine took revenge by crying (*Conf.* 1.6.8). The gift of the yet inarticulate infant to communicate his emotions nonverbally in these forms is viewed by Augustine as a peculiarity of man which has been granted by God (*Conf.* 1.6.10–1.7.12). Next he reports how he acquired the verbal expressions for things:

> *cum ipsi appellabant rem aliquam et cum secundum eam vocem corpus ad aliquid movebant, videbam et tenebam hoc ab eis vocari rem illam quod sonabant cum eam vellent ostendere. hoc autem eos velle ex motu corporis aperiebatur tamquam verbis naturalibus omnium gentium, quae fiunt vultu et nutu oculorum ceterorumque membrorum actu et sonitu vocis indicante affectionem animi in petendis, habendis, reiciendis fugiendisve rebus.* (*Conf.* 1.8.13)

> 'When they named anything, and when at that name they moved their bodies toward that thing, I observed it and gathered thereby that that word which they then pronounced was the very name of the thing which they showed me. And that they meant this or that thing, was discovered to me by the motion of their bodies, even by that natural language, as it were, of all nations, which expressed by the countenance and cast of the eye, by the action of other parts, and the sound of the voice, discovers the affections of the mind, either to desire, enjoy, refuse, or to avoid anything.'

A similar point can be found in Lucretius' theory of the beginnings of civilization, expounded in the fifth book of his didactic poem *De rerum natura*[12]. He is convinced that the invention of language was not the work of an individual because it is not plausible that this ability should have been given only to a single human being. Moreover, how should such a 'first inventor' have taught his created expressions to others (*De rer. nat.* 5.1041–55)? Instead, one has to assume that in the beginning man uttered various sounds for different impressions and feelings and that, in doing so, he contrasted the things he wished to denote (*De rer. nat.* 5.1056–90). In its initial stage, human language was comparable to the articulation of infants, who, to illustrate their messages, fall back upon gestures and point to things they refer to, using their fingers (*De rer. nat.* 5.1028–33, cf. already 5.1021–23).

Despite many attempts in ancient literature at comprehending elements of nonverbal communication as a medium of international communication, the ancients never attempted to construct a verbal equivalent, an international auxiliary language or planned language like the modern Esperanto or Volapük[13]. This fact is easily explained: On the one hand, Greek and, later, Latin have the status of world languages in the entire Mediterranean world, not only by the self-definition of their respective native speakers but as a matter of fact. On the other hand, it would have been completely unthinkable for a Greek or a Roman to construct an artificial language as a replacement for a historically evolved idiom simply for the purpose of communication with other nations; it was the pronounced ethno- and linguocentrism that militated against such an idea right from the beginning[14]. Moreover, well-directed attempts at surmounting otherness and the striving for intercultural exchange enhanced by the learning of foreign languages are alien to ancient thinking, even if the idea of cosmopolitism and the equality of men found its supporters in some philosophical schools, notably the Stoics and Cynics[15]. Still, it has to be borne in mind that, at least for the Roman world, the naïve opinion that the structure of the mother tongue is given by nature and the only one which is 'correct'—a view which equates 'mother tongue' and 'language'—was not disseminated in the literary sources we have. This is not least a result of the Greco-Latin bilingualism in the Roman upper-class which had been increasingly noticeable since the second century B.C.

2. GESTURE, FACIAL EXPRESSION AND VOICE IN RHETORIC. Theoretical aspects of nonverbal communication are most extensively dealt with by ancient authors in rhetorical treatises as prescriptive and normative instructions for the orator to enable him to

be as persuasive as possible over and above what is achieved by linguistic and stylistic means. It is evident that, among the orator's five tasks, gesture, facial expression and voice quality form part of delivery (ὑπόκρισις resp. *actio* or *pronuntiatio*). The important role delivery played among the orator's tasks (*officia oratoris*) can be gathered from an anecdote about Demosthenes which is handed down by many Greek and Roman authors: He is said to have attributed first, second and third rank to delivery[16], and in this he is followed by later theoreticians (see e.g., *Rhet. Her.* 3.19; Cicero, *De orat.* 3.213). Most Greek writings on gesture, facial expression and voice quality have unfortunately been lost[17] so that in the majority of cases we have to content ourselves with reconstructing Greek theory from Roman sources, not all of which have survived either[18]. In the following, I shall concentrate on passages from four Latin works, singling out no more than the most important aspects: the anonymous *Rhetorica ad Herennium* (3.19–27), Cicero's *Orator* (54–60) and *De oratore* (esp. 3.213–27), and Quintilian's *Institutio oratoria* (esp. 11.3). All these writers divide delivery into voice quality (*vox*) on the one hand, and gesture (*gestus*) and facial expression (*vultus*) on the other. Accordingly, delivery was termed either *pronuntiatio* with emphasis on the vocal aspects or *actio* with reference to the use of body language (*Rhet. Her.* 3.19; Cicero, *Orator* 55 and *De orat.* 3.213–27; Quintilian, *Inst. orat.* 11.3.1, 11.3.14).

With the exception of Quintilian, the writers surveyed pay much more attention to the use of the voice than to the use of the body. Since by the different modulations of his voice the orator evokes certain emotions in his audience, he has to adjust his voice to the contents of his speech and to the reactions he wants to elicit from his listeners. This intention requires the chosen pitch to be in harmony with the orator's character (particularly remarkable *Rhet. Her.* 3.22). The state of mind the voice evokes in the audience ought to be reinforced by corresponding gestures and facial expressions of the orator. The principle of appropriateness (*aptum*) is equally vital for the use of elements of nonverbal communication as it is on the verbal level, where the choice of words, the stylistic levels and registers must be selected in accordance with the topic of the orator's speech, his intention, the make-up of his audience and other external circumstances, but also with his own personal make-up[19]. Elements of gesture and facial expression should fit the orator's individual personality and at the same time match the requirements of the oratorical profession, whose representatives are expected to be concerned about their dignity (*dignitas*) and authority (*gravitas*). For that reason, the orator should not behave like an actor on stage, even if dramatic art may sometimes provide the orator with ideas for the shaping of his speeches especially at the nonverbal level (*Rhet. Her.* 3.26; Cicero, *De orat.* 3.220; Quintilian, Inst. *orat.* 11.3.103, 123, 125, 181–84, esp. 184 on the situation in Quintilian's era). In connection with this, Quintilian recommends a restriction to gestures which accompany speech in a natural way ('illustrators' in the terminology of Ekman and Friesen) and suggests that gestures that substitute the spoken word ('emblems' in Ekman's and Friesen's terms; cf. 1969: 63–70, and 1979: 111ff., 118) should be avoided because they belong to the domain of actors in the theatre (*Inst. orat.* 11.3.88–91). The reason for such a demand is apparent, if we consider the social stratification of Roman society.

Actors and pantomimes usually belonged to the lower stratum; their profession was associated with ill repute, a lack of moral restraint and even indecency which at least officially conflicted with the strict morality of respectable Roman citizens (cf. Wüst 1949: esp. 860–63; Rotolo 1957: 49f., 63–67).

In rhetorical writings, the significance of facial expression for the orator and in particular the expression of the eyes are firmly emphasized. For example, Cicero remarks[20]:

> *Sed in ore sunt omnia, in eo autem ipsi dominatus est omnis oculorum; …animi est enim omnis actio et imago animi vultus, indices oculi: nam haec est una pars corporis, quae, quot animi motus sunt, tot significationes et commutationes possit efficere; neque vero est quisquam, qui eadem conivens efficiat.* (*De orat.* 3.221)
>
> 'But everything depends on the face, while the face itself is entirely dominated by the eyes… For delivery is wholly the concern of the feelings, and these are mirrored by the face and expressed by the eyes; for this is the only part of the body capable of producing as many indications and variations as there are emotions, and there is nobody who can produce the same effect with his eyes shut.'

As regards gestures, Quintilian believes that the hands are most important, since with them the orator can transmit messages in a way that is nearly as sophisticated as the use of words. In denoting places and persons, the hands take on, as it were, the function of adverbs and pronouns. Since with them one can describe an incredible range of facts, Quintilian thinks that it is no exaggeration to understand gestural communication which makes use of the hands as a language common to all human beings (*Inst. orat.* 11.3.85–87).

3. CONCLUSION. I do not want to conclude my overview without mentioning that I have presented a very selective treatment of the subject. Among the ancient treatises on rhetoric, for example, particularly the eleventh book of Quintilian's *Institutio oratoria* would be worth a separate analysis (cf. as a first start Maier-Eichhorn 1989), since it is the only extant comprehensive treatment of the use of gestures in speeches. In addition to such large-scale treatises, there are numerous ancient texts which contain some valuable information on the more theoretical aspects of the use of gestures and facial expression. Let me just refer to the famous passage in Plato's *Cratylus* (esp. 422 D11–423 B7) or to the Younger Seneca's prose writings.

But even without this, I hope to have demonstrated that the theorizing on nonverbal communication by ancient authors starts from various perspectives. The basis of these reflections are, first, hypotheses on the origin of language; second, considerations about the first-language acquisition of the child; third, remarks on the problem of foreign languages and, connected with this, the view of gesture and facial expression as a universal language of all human beings; and fourth, the elaboration of rhetorical doctrines for the purpose of a skilful use of nonverbal elements in public speech. It is evident that whatever perspective the different writings may take, they also rely upon sign theory and discuss the status of signs. A systematic compilation of theoretical reflections on nonverbal communication in ancient Greek and Roman authors is still a desideratum; it is the aim of my latest research project to fill this gap in a more extensive study.

---

[1] On the interdisciplinary framework of scholarship on nonverbal communication e.g., Noll (1998:156).

[2] On the history of theorizing on gesture see the most recent overview in Müller (1998: 25–85); the author includes antiquity, in particular Quintilian.

[3] The most exhaustive study is still Sittl (1890), most recent monographs by Aldrete (1999) and Boegehold (1999). Shorter accounts can be found in Wöhrle (1990), Bremmer (1991), Graf (1991), and Lateiner (^{3}1996). The following studies are concentrated on single ancient authors: Lateiner (1987), Lateiner (1995), Maier-Eichhorn (1989), and Lobe (2000). Only very briefly on antiquity the article by Košenina (1996: esp. 565–67).

[4] See in particular Déonna (1914), Zschietzschmann (1924), Jucker (1956), Brilliant (1963), Neumann (1965), Bogen (1969), and Fehr (1979).

[5] Regarding Rome on the political level as the model for everything else, the Romans were, on the whole, not very open-minded about anything non-Roman. A knowledge of foreign languages extending beyond Greek is not very often a topic of Latin writers; sometimes the ability to manage one foreign language is already regarded as a remarkable achievement. Important sources are e.g., Quintilian, *Inst. orat.* 11.2.50: *Ceterum quantum natura studioque valeat memoria, vel Themistocles testis, quem unum intra annum optime locutum esse Persice constat, vel Mithridates, cui duas et viginti linguas, quot nationibus imperabat, traditur notas fuisse, vel Crassus ille dives, qui cum Asiae praeesset, quinque Graeci sermonis differentias sic tenuit, ut, qua quisque apud eum lingua postulasset, eadem ius sibi redditum ferret… ('For the rest there are many historical examples of the power to which memory may be developed by natural aptitude and application. Themistocles is said to have spoken excellently in Persian after a year's study; Mithridates is recorded to have known twenty-two languages, that being the number of the different nations included in his empire; Crassus, surnamed the Rich, when commanding in Asia had such a complete mastery of five different Greek dialects, that he would give judgement in the dialect employed by the plaintiff in putting forward his suit')*; on Themistocles cf. already Thucydides 1.138 and Nepos, *Them.* 10.1, on Mithridates similarly Pliny, *Nat. hist.* 7.88, 25.6, and Gellius, *Noct. Att.* 17.17.2, and on P. Licinius Crassus Dives cf. Valerius Maximus 8.7.6. Also relevant is Gellius, *Noct. Att.* 17.17.1: *Quintus Ennius tria corda habere sese dicebat, quod loqui Graece et Osce et Latine sciret. ('Quintus Ennius used to say that he had three hearts, because he knew how to speak Greek, Oscan, and Latin.')*

[6] On the problem of foreign languages in Greco-Roman antiquity see Fögen (2000: esp. 31–60), with all relevant secondary literature (see also pp. 242–47).

[7] Cf. also Weinrich (1985:198): 'Wahrscheinlich ist die Rat- und Hilflosigkeit gegenüber einem anderen Menschen, mit dem man sich nicht verständigen kann, eine menschliche Grund- und Grenzerfahrung, bei der Andersheit und Fremdheit zusammenfallen, wenigstens auf den ersten Blick.'

[8] The importance of the *societas patrii sermonis* is similarly emphasized in Anonymus, *Rhet. Her*. 4.37: *nihil Fregellanis morum et sermonis societas opitulata est ('of no avail, to the Fregellans was fellowship with us in customs and in language')*, and in Cicero, *Verr.* 2.5.167 (taken up by Aulus Gellius, *Noct. Att.* 1.7.2): *…apud cives… Romanos, qui et sermonis et iuris et multarum rerum societate iuncti sunt ('…among Roman citizens, who are united by the common bond of language, rights, and many interests')*, and later on Charisius, *GL* 1.50.15–16 Keil (= 62.2–4 Barwick): *Latinus vero sermo cum ipso homine civitatis suae natus significandis intellegundisque quae diceret praestitit ('But the Latin language, come into existence at the same time as the Roman citizen, succeeded in expressing its message by conveying meaning and having it understood')*. That such a view was not restricted to the Romans, can be gathered from Curtius Rufus, *Hist. Alex. Mag*. 6.10.23, who lets the Macedonian Philotas point out: *Mihi quidem obicitur quod societatem patrii sermonis asperner, quod Macedonum mores fastidiam ('It is even charged against me that I scorn association with my native language, that I disdain the customs of the Macedonians')*. Cf. also the viceversa relation in Quintilian, *Inst. orat.* 8.1.3: *…si fieri potest, et verba omnia et vox huius alumnum urbis oleant, ut oratio Romana plane videatur, non civitate donata ('…if possible, our voice and all our words should be such as to reveal the native of this city, so that our speech may seem to be of genuine Roman origin, and not merely to have been presented with Roman citizenship.')*. Language as a vital power, which re-enforces the sense of community, in the following case of the Athenians as belonging to the larger body of the Hellenes, is—apart from common religious customs and origin (ties of blood)—already mentioned in Herodotus 8.144 (cf. also 2.56f.). For the later period see Dionysius Halicarnassus, *Antiqu. Rom.* 1.89.3.

[9] An excellent summary of the different positions from universalism to relativism can be found in Wallbott (1979) and Wallbott (1993); see also the short remarks in Clyne (1975:162–63). Supporters of a moderate universalism include Ekman (1970) and Ekman and Friesen (1979:111–12, 118). The universalist hypothesis was first and foremost rejected by Birdwhistell (1970, esp. 34).

[10] On the problem of the authorship and date see Robertson (1913): the work was written between 162 and 165 during Lucian's stay in Antiochia; with it, the author wanted to win over the emperor Verus, who was enthusiastic about pantomimes (Script. Hist. Aug., *Verus* 8 and 10–11.). This view is shared by Robert (1930, esp. 120–22); in contrast Bier (1917) suggests a date between Domitian and Marcus Aurelius. That the work is not by Lucian is assumed by Helm (1906:369f). On parallels in Aristides and Libanius cf. Mesk (1908:59–74). A useful study on the ancient pantomime dance in general is, apart from Bier (1917), Rotolo (1957) who also collects relevant epigraphic and epigrammatic sources.

[11] See the commentaries by O'Donnell (1992) and Clark (1995). Also important is Marrou (1981).

[12] On language origin in Epicurus (*Ep. ad Herod.* 75–76) and Lucretius see the following studies: de Lacy (1939), Long (1971), Offermann (1972), Schrijvers (1974), Pigeaud (1983), Ferguson (1987), Pfligersdorffer (1988), and Hossenfelder (1991).

[13] Cf. Schwyzer (1902) and Rochette (1995:10–11).

[14] Cf. also Pei (1958:83): 'Universality was implicit in the Graeco-Roman world. To its members that world constituted the universe, despite the vague awareness of a world beyond, …The Roman Empire, once constituted, became a powerful, living reality whose limes was almost as much of a boundary to the ancients as the stratosphere is to us.'

[15] On this see Baldry (1965) and Lana (1951).

[16] Cicero, *Brutus* 142, *Orator* 56, *De orat.* 3.213; Valerius Maximus 8.10 ext. 1; Quintilian, *Inst. orat.* 11.3.6–7; Philodemus, *Rhet.* 1.196 Sudhaus; Longinus, *Rhet. Gr.* 1.2.195 Spengel/Hammer; Plutarch, Mor. 845B. See also Dionysius Halicarnassus' remarks on Demosthenes' intensive voice and body exercises in *Dem.* 53.

[17] Brief treatments of this are Maier-Eichhorn (1989:7–14) and Katsouris (1989:26–33). On Theophrastus see Fortenbaugh (1985). The non-existence of more extensive writings on *actio* is mentioned in *Rhet. Her.* 3.19 (cf. also 3.27 fin.).

[18] E.g., Quintilian, *Inst. orat.* 11.3.143, refers to writings on gesture by Plotius Gallus and Nigidius Figulus.

[19] Cf. Fögen (1999:5–9), similarly Fögen (2000:119–22). In addition, see Quintilian, *Inst. orat.* 11.1, 11.3.61-65, 11.3.150–53, and further 11.3.177–84, esp. 180.

[20] Similarly Quintilian, Inst. orat. 11.3.75–77, esp. 75: *sed in ipso vultu plurimum valent oculi, per quos maxime animus emanat, ut citra motum quoque et hilaritate enitescant et tristitiae quoddam nubilum ducant. quin etiam lacrimas iis natura mentis indices dedit, quae aut erumpunt dolore aut laetitia manant. motu vero intenti, remissi, superbi, torvi, mites, asperi fiunt: quae, ut actus poposcerit, fingentur. ('But of the various elements that go to form the expression, the eyes are the most important, since they, more than anything else, reveal the temper of the mind, and without actual movement will twinkle with merriment or be clouded with grief. And further, nature has given them tears to serve as interpreters of our feelings, tears that will break forth for sorrow or stream for very joy. But, when the eyes move, they become intent, indifferent, proud, fierce, mild, or angry; and they will assume all these characters according as the pleading may demand.')*

SELECTED BIBLOGRAPHY

(EDITORS' NOTE: This section contains references to background material not cited in the paper.)

ACHARD, GUY. 1991. *La communication à Rome*. Paris: Les Belles Lettres.

ALDRETE, GREGORY S. 1999. *Gestures and acclamations in ancient Rome*. Baltimore: Johns Hopkins University Press.

BALDRY, H. C. 1965. *The unity of mankind in Greek thought*. Cambridge: Cambridge University Press.

BARNETT, DENE. 1996. Headword 'Gestik'. In *Historisches Wörterbuch der Rhetorik*, vol. 3, ed. by Gert Ueding, 972–89. Darmstadt: Wissenschaftliche Buchgesellschaft.

BIER, HUGO. 1917. *De saltatione pantomimorum.* Diss. Bonn.

BIRDWHISTELL, RAY L. 1970. *Kinesics and context.* Philadelphia: University of Pennsylvania Press.

BOEGEHOLD, ALAN L. 1999. *When a gesture was expected: A selection of examples from archaic and classical Greek literature*. Princeton: Princeton University Press.

BOGEN, KATHRIN. 1969. *Gesten in Begrüßungsszenen auf attischen Vasen*. Bonn: Habelt.

BREMMER, JAN. 1992. Walking, standing, and sitting in ancient Greek culture. In Bremmer & Roodenburg 1992:15–35.

——— & HERMAN ROODENBURG (eds.). 1992. *A cultural history of gesture*. Ithaca NY: Cornell University Press.

BRILLIANT, RICHARD. 1963. *Gesture and rank in Roman art: The use of gestures to denote status in Roman sculpture and coinage*. New Haven: Connecticut Academy of Arts & Sciences.

CLARK, GILLIAN. 1995. *Augustine – Confessions, Books I–IV.* Cambridge: Cambridge University Press.

CLYNE, MICHAEL. 1975. *Forschungsbericht Sprachkontakt. Untersuchungsergebnisse und praktische Probleme*. Kronberg: Scriptor.

DÉONNA, WALDEMAR. 1914. *L'expression des sentiments dans l'art grec.* Paris: Laurens.

EKMAN, PAUL. 1970. Zur kulturellen Universalität des emotionalen Gesichtsausdrucks. In Scherer & Wallbott 1979:50–58.

——— & WALLACE V. FRIESEN. 1969. The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica* 1:49–98.

——— & ———. 1979. Handbewegungen. In Scherer & Wallbott 1979:111–18.

EVANS, ELIZABETH C. 1935. Roman descriptions of personal appearance in history and biography. *Harvard studies in classical philology* 46:43–84.

———. 1969. Physiognomics in the ancient world. *Transactions of the American Philosophical Society N.S.* 59.5:3–101.

FANTHAM, ELAINE. 1982. Quintilian on performance: Traditional and personal elements in *Institutio* 11.3. *Phoenix* 36:243–63.

FEHR, BURKHARD. 1979. *Bewegungsweisen und Verhaltensideale: Physiognomische Deutungsmöglichkeiten der Bewegungsdarstellung an griechischen Statuen des 5. und 4. Jh. v. Chr.* Bad Bramstedt: Moreland.

FERGUSON, JOHN. 1987. Epicurean language theory and Lucretian practice. *Liverpool classical monthly* 12:100–5.

FÖGEN, THORSTEN. 1999. Spracheinstellungen und Sprachnormbewußtsein bei Cicero. *Glotta* 75:1–33.

———. 2000. *'Patrii sermonis egestas': Einstellungen lateinischer Autoren zu ihrer Muttersprache. Ein Beitrag zum Sprachbewußtsein in der römischen Antike*. München: Saur.

FORTENBAUGH, WILLIAM W. 1985. Theophrastus on delivery. In *Theophrastus of Eresus: On his life and work*, ed. by William W. Fortenbaugh, Pamela M. Huby & Anthony A. Long, 269–88. New Brunswick: Transaction Books.

———. 1986. Aristotle's Platonic attitude toward delivery. *Philosophy and rhetoric* 19:242–54.

GRAF, FRITZ. 1992. Gestures and conventions: The gestures of Roman actors and orators. In Bremmer & Roodenburg 1992:36–58.

GRAJEW, FELIX. 1934. *Untersuchungen über die Bedeutung der Gebärden in der griechischen Epik* (Diss. Freiburg im Breisgau 1930). Berlin.

HELM, RUDOLF. 1906. *Lucian und Menipp*. Leipzig:Teubner.

HOLOKA, JAMES P. 1992. Nonverbal communication in the classics: research opportunities. In *Advances in nonverbal communication: Sociocultural, clinical, esthetic* *and literary perspectives*, ed. by Fernando Poyatos, 237–54. Amsterdam: John Benjamins.

Hossenfelder, Malte. Epikureer. In *Geschichte der Sprachtheorie. Band 2: Sprachtheorien der abendländischen Antike*, ed. by Peter Schmitter, 217–37. Tübingen: Narr.

Hurschmann, Rolf. 1998. Headword 'Gebärden. iii. Griechenland und Rom'. In *Der Neue Pauly (Band 4)*, 819–28. Stuttgart: Metzler.

Jucker, Ines. 1956. *Der Gestus des Aposkopein. Ein Beitrag zur Gebärdensprache in der antiken Kunst*. Zürich: Juris.

Katsouris, Andreas G. 1989. *Rhetorike Hypokrise.* Ioannina: Panepistemio Ioanninon.

Kayser, Johannes. 1910. Theophrast und Eustathius περὶ ὑποκρίσεως. *Philologus* 69:327–58.

Key, Mary Ritchie. 1992. Headword 'Nonverbal communication'. In *International encyclopedia of linguistics* (vol. 3), ed. by William Bright, 107–10. New York: Oxford University Press.

Kötting, Bernhard. 1978. Headword 'Geste und Gebärde'. *Reallexikon für Antike und Christentum* 10:895–902.

Košenina, Alexander. 1996. Headword 'Gebärde'. In *Historisches Wörterbuch der Rhetorik* (vol. 3), ed. by Gert Ueding, 564–79. Darmstadt: Wissenschaftliche Buchgesellschaft.

de Lacy, Philip H. 1939. The Epicurean analysis of language. *American journal of philology* 60:85–92.

Lana, Italo. 1951. Tracce di dottrine cosmopolitiche in Grecia prima del cinismo. *Rivista di filologia e di istruzione classica* 29:193–216, 317–38.

Lateiner, Donald. 1987. Nonverbal communication in the *Histories* of Herodotus. *Arethusa* 20:83–119.

———. 1995. *Sardonic smile: Nonverbal behavior in Homeric epic*. Ann Arbor: University of Michigan Press.

———. ^{3}1996. Headword 'Gestures'. In *The Oxford Classical Dictionary*, ed. by Simon Hornblower & Antony Spawforth, 635–36. Oxford: Oxford University Press.

Lobe, Michael. 2000. *Die Gebärden in Vergils Aeneis. Zur Bedeutung und Funktion von Körpersprache im römischen Epos*. Frankfurt am Main: Lang.

Long, Anthony A. 1971. Aisthesis, prolepsis and linguistic theory in Epicurus. *Bulletin of the Institute of Classical Studies* 18:114–33.

Maier-Eichhorn, Ursula. 1989. *Die Gestikulation in Quintilians Rhetorik.* Frankfurt am Main: Lang.

Marrou, Henri-Irénée. 1981. *Augustinus und das Ende der antiken Bildung*. Paderborn: Schöningh. [Translation of *Saint Augustin et la fin de la culture antique*, Paris ^{4}1958, by Lore Wirth-Poelchau & Willy Geerlings, revised by Johannes Götte]

Mesk, Josef. 1908. Des Aelius Aristides verlorene Rede gegen die Tänzer. *Wiener Studien* 30:59–74.

Müller, Cornelia. 1998. *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Spitz.

Neumann, Gerhard. 1965. *Gesten und Gebärden in der griechischen Kunst*. Berlin: de Gruyter.

Noll, Thomas. 1998. Wie blättert man in einem Gestikon? In *Körperbewegungen und ihre Bedeutungen*, ed. by Caroline Schmauser & Thomas Noll, 156–72. Berlin: Spitz.

O'Donnell, James J. 1992. *Augustine – Confessions* (3 vol.). Oxford: Oxford University Press.

Offermann, Helmut. 1972. Lukrez V 1028–1090. *Rheinisches Museum* 115:150–56.

Pei, Mario. 1958. *One language for the world*. New York: Devin-Adair.

Pfligersdorffer, Georg. 1988. Zur Sprachentstehung nach Lukrez. In *Antike Rechts- und Sozialphilosophie*, ed. by Olof Gigon & Michael W. Fischer, 138–46. Frankfurt am Main: Lang.

Pigeaud, J. 1983. Épicure et Lucrèce et l'origine du langage. *Revue des Études Latines* 61:122–44.

Robert, Louis. 1930. Pantomimen im griechischen Orient. *Hermes* 65:106–22.

Robertson, D. S. 1913. The authenticity and date of Lucian 'De saltatione'. In *Essays and studies presented to William Ridgeway*, ed. by Edmund C. Quiggin, 180–85. Cambridge: Cambridge University Press.

Rochette, Bruno. 1995. Grecs et Latins face aux langues étrangères. Une contribution à l'étude de la diversité linguistique dans l'antiquité classique. *Revue Belge de Philologie et d'Histoire* 73:5–16.

Rotolo, Vincenzo. 1957. *Il pantomimo. Studi e testi*. Palermo: Presso L'Accademia.

———. 1972. La communicazione linguistica fra alloglotti nell'antichità classica. In *Studi classici in onore di Quintino Cataudella*, vol. 1, 395–414. Catania: Facoltà di Lettere e Filosofia.

Scherer, Klaus R. 1970. *Non-verbale Kommunikation. Ansätze zur Beobachtung und Analyse der außersprachlichen Aspekte von Interaktionsverhalten*. Hamburg: Buske.

——— & Harald G. Wallbott (eds.). 1979. *Nonverbale Kommunikation: Forschungsberichte zum Interaktionsverhalten*. Weinheim: Beltz.

Schottlaender, Rudolf. 1967. Der römische Redner und sein Publikum. *Wiener Studien* 80:125–46.

Schrijvers, P. H. 1974. La pensée de Lucrèce sur l'origine du langage (*DRN* V 1019–1090). *Mnemosyne* 4.27:337–64.

Schwyzer, Eduard. 1902. *Die Weltsprachen des Altertums in ihrer geschichtlichen Stellung*. Berlin: Weidmann.

Sittl, Carl. 1890. *Die Gebärden der Griechen und Römer*. Leipzig: Teubner.

Solmsen, Friedrich. 1938. Aristotle and Cicero on the orator's playing upon the feeling. *Classical philology* 33:390–404.

Sonkowsky, Robert P. 1959. An aspect of delivery in ancient rhetorical theory. *Transactions and proceedings of the American Philological Association* 90:256–74.

TALADOIRE, BARTHÉLÉMY-ANTONIN. 1951. *Commentaires sur la mimique et l'expression corporelle du comédien romain.* Montpellier: Faculté des Lettres de l'Université de Montpellier.

WALLBOTT, HARALD G. 1979. Gesichtsausdruck. In Scherer & Wallbott 1979:35–42.

———. 1993. Mimik und Emotion: Anmerkungen aus Sicht der neueren nonverbalen Kommunikationsforschung. In *Körpersprache. Diagnostik und Therapie von Sprach-, Sprech- und Stimmstörungen*, ed. by Geert Lotzmann, 26–41. München: Reinhardt.

WARNECKE, BORIS. 1910. Gebärdenspiel und Mimik der römischen Schauspieler. *Neue Jahrbücher für das klassische Altertum* 25:580–94.

WEINRICH, HARALD. 1985. Fremdsprachen als fremde Sprachen. In *Wege der Sprachkultur*, by Harald Weinrich, 198ff. Stuttgart: Deutsche Verlagsanstalt.

WÖHRLE, GEORG. 1990. Das fünfte officium des antiken Redners. *Gymnasium* 97:31–46.

WÜLFING, PETER. 1995. Antike und moderne Redegestik. Eine frühe Theorie der Körpersprache bei Quintilian. In *Kommunikation durch Zeichen und Wort. Stätten und Formen der Kommunikation im Altertum IV*, ed. by Gerhard Binder & Konrad Ehlich, 71–90. Trier: Wissenschaftlicher Verlag.

———. 1998. Headword 'Gestus'. In *Der Neue Pauly (Band 4)*, 1021–24 Stuttgart: Metzler.

WÜST, ERNST. 1949. Pantomimus. In *RE* XVIII.3, 833–69.

ZSCHIETZSCHMANN, WILLY. 1924. *Untersuchungen zur Gebärdensprache in der älteren griechischen Kunst*. Diss. Jena.

* * *

THE GRAMMAR OF THE PICTISH SYMBOL STONES

Toby D. Griffen
Southern Illinois University Edwardsville

THE PICTS inhabited most of Scotland in the first millennium of the Common Era. While we know from such evidence as place names (Nicolaisen 1972, 1996) that they spoke a Brythonic Celtic language, aside from these place names, the king lists (see Cummins 1995), and a handful of ogham inscriptions that have so far eluded decoding (but see Cummins 1999:60–68), nothing else is known of their language (compare Forsyth 1996, 1997a).

1. THE PICTISH SYMBOL STONES. The most extensive and most identifiably Pictish artifacts that these people left behind are the symbols that they carved on a number of stones and other articles. These display a remarkable degree of consistency and are reproduced in typical forms in Table 1 (overleaf, from the compilation of Sutherland 1997, who draws her figures from Allen & Anderson 1903).

Where the interests of linguistics enter the picture is in the enigmatic combinations of these symbols throughout Pictland. Usually in pairs, one over another, these combinations strongly suggest the communication of some form of information adhering to a set of rules. As Thomas points out:

> Statements containing two symbols, or Pairs, do not appear to conform to any random distribution (this could almost certainly be demonstrated mathematically by a simple matrix analysis) and it may be assumed that selective rules are at work. There is a noticeable emphasis on the use of Object, rather than Animal, symbols. Repetition of complete pairs exists, and among the actual symbols used, both Animal and Object, certain symbols occur disproportionately more often than others. (Thomas 1963:37)

Such observations indicate to the linguist that some form of rule-based grammar is involved; and one mathematical matrix has been worked out by Jackson (1984), although with an interpretation that is not widely accepted (Forsyth 1997b:85). Of course, before we attempt to ascertain what they might have meant, we must first determine the level of grammar on which this information was being conveyed.

1.1 THE SEMANTIC LEVEL. The first (and perhaps the most obvious) level suggested for the grammar of the Pictish symbol stones is the semantic, or more precisely the lexico-semantic. In this approach, each symbol relates to some precise designation that maintains a particular meaning.

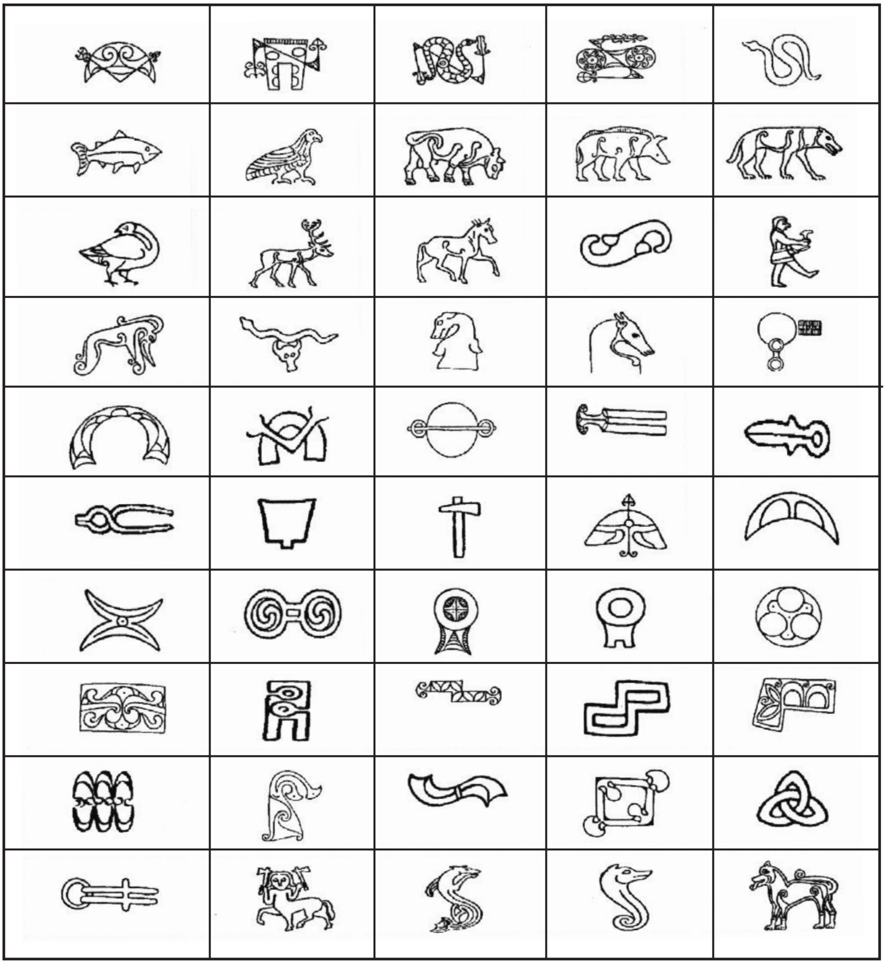


Table 1. Pictish symbols.

Thomas (1963) interprets the symbols as funerary. Since there are not enough such symbols to represent individuals or indeed individual names, he suggests that they 'must perforce do this in terms of… status and… group-affiliation' (Thomas 1963:87). Moreover, he posits quasi-grammatical rules to account for their distribution, considering them to be personal nouns and adjectives.

Henderson (1971), while accepting the semantic and even the grammatical basis for the symbols, suggests that they may have referred rather to the ownership of land. Along somewhat similar lines, Jackson (1984) expands upon the grammatical basis and presents an impressive hypothesis that the paired symbols represent marriages within the Pictish matrilineal system. Unfortunately, the hypothesis suffers from some rather serious problems (Driscoll 1986).

More recently, Samson (1992) proffers the hypothesis that the symbols alone reveal nothing—only the symbols in combination have significance. Each combination supposedly refers to an individual memorialized on the stones. Indeed, such pairing of thematic elements does reflect the construction of names used in contemporary Irish and Anglo-Saxon, and a statistical analysis of such elements appears to support the hypothesis. The drawback to this approach, however, lies in the fact that these symbols do occasionally occur singly; nor are they restricted to funerary contexts, but are even found 'on a number of portable objects, including a silver pin and silver plaques from Norrie's Law, Fife, and the terminal of the silver chain from Whitecleuch, Lanarkshire' (Ritchie & Ritchie 1981:164; see also 172–74).

1.2 THE ORTHOGRAPHIC LEVEL. One significant breakthrough in Samson's argument is that it takes the debate off of the semantic level *per se* and onto a morphological one. As noted by Forsyth, 'for the first time… the symbols represented **language** rather than **ideas**' (1997b:87). This suggests to her that the symbols might reside on a level of writing—an orthographic level.

In examining how such a system might work, however, Forsyth has to acknowledge serious problems for each type of known orthographic system. First of all, a logographic system would not work because there are simply not enough symbols. While the number of symbols would certainly support a syllabary, Pictish words and names were not restricted to two syllables each. Likewise, although the number of symbols could support an alphabet, the 'brevity of the extant texts . . . all but precludes an alphabetic interpretation for Pictish symbols (unless one is prepared to accept the unlikely scenario that the symbol statements represent pairs of initials)' (Forsyth 1997b:93).

2. AN APPEAL TO CONTEXT. The greatest hindrance to a linguistic interpretation of the symbol stones has been a lack of appropriate context. Harkening back to linguistics in the early 20th century, we recall that in order for communication to be comprehended and analyzed effectively, we first need to know the 'context of situation.' As noted by Firth,

> The basic assumption of the theory of analysis by levels is that any text can be regarded as a constituent of a *context of situation* or of a series of such contexts, and thus attested in experience, since the categories of the abstract context of situation will comprise both verbal and non-verbal constituents and, in renewal of connection, should be related to an observable and justifiable grouped set of events in the run of experience. (Firth 1968:175)

Indeed, there has been some recent evidence that may shed light upon the intended context of situation for these symbols, and hence upon the appropriate level of grammar for analysis.

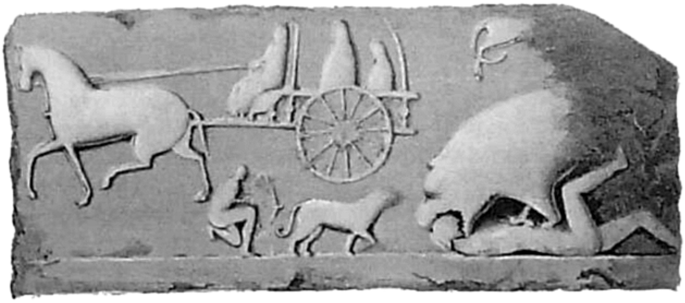


***Figure 1**. Meigle 10.*

2.1 THE CULT OF THE ARCHER GUARDIAN. While a full explanation of the context for the Cult of the Archer Guardian is not possible in this forum, a brief summary of some of the evidence may suffice. In a recumbent stone known as Meigle 10 (following the enumeration of Allen and Anderson 1903), we find a situation depicted by Chalmers (1848) as in Figure 1 (from Ritchie 1997:121).

As explained with detailed justification elsewhere (see Griffen 2000), the figures being transported in the formal processional vehicle—a *carpentum* (see Laing & Laing 1984:277–78)—are souls undergoing the translation to the otherworld according to widely held beliefs in the Pagan Celtic religion (compare, for example, Green 1986:121–37). The scene in the lower right shows the alternative fate of the soul—its capture by what we may term the Decapitating Beast through the removal of the person's head. In the Celtic tradition, the head is the home of the soul, and both the role of the head and its decapitation are also well attested in the Celtic tradition (Ross 1996:94–95). In between these two scenes, we see the Archer Guardian, who protects the otherworld and those being translated from the ravages of the Decapitating Beast.

In the Christianization of Pictland, the Archer Guardian was taken over as the one who guards heaven for the elect, at least according to the doctrine of predestination, which was in fact prevalent in Britain at the time (see Hardinge 1972:61). We find rather clear evidence of this syncretism in the Ruthwell Cross in Figure 2 (detail from Meyvaert 1992: figure 4), in which the Archer Guardian is shown in the lower panel of the cross, metaphorically between heaven and hell. The Celtic sacred eagle, appropriately enough, is in the upper panel, for its 'vast wing-span and powerful, high flight epitomize the huge span of the sky' (Green 1986:188).

2.2 THE V-ROD AND THE Z-ROD. When we examine the symbols in table 1, we see that there are two types that stand out—those with the so-called V-rod and Z-rod (reproduced at the beginning of the table for ease of reference). The V-rod is quite clearly an arrow that has been bent to an angle of about 90° and placed over a crescent. The crescent is widely regarded as a symbol of death (compare Sutherland 1994:103-107), and the arrow is bent in precisely the manner that one would bend a sword or other such votive offering before depositing it—a ritual 'killing' of the object (compare Green

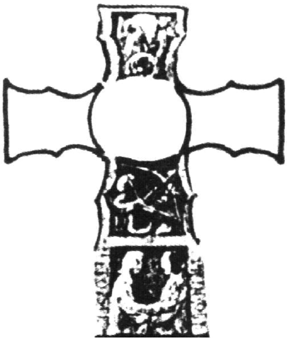


Figure 2. *The Ruthwell cross.*

1995:470-71, Maier 1997:241). This very widespread combination—the only one with a V-rod—is thus quite consistent with the cult of the Archer Guardian and probably represents a symbolic invocation of the Archer perhaps upon the death of an individual.

The Z-rod occurs superimposed upon only three symbols (the second through the fourth symbols, respectively, in table 1): The 'notched rectangle' is widely regarded as a chariot, and as such it is completely consistent with the carpentum in which the soul is translated to the otherworld or to heaven (compare Thomas 1963:52-53); the snake is a particularly holy creature in the Pagan Celtic religion because it travels between this world and the otherworld (in its underworld aspect—Sutherland 1994:89-90); and the double-disk symbol is a double-sun representation that in its broadest interpretation can certainly represent this world and the otherworld and the path in between (compare Sutherland 1994:107-108).

When we examine the Z-rod in these contexts, we find that they are all very much consistent with the cult of the Archer Guardian. The rod itself is an arrow (or spear) bent twice so as to define the boundary between this world and the otherworld—the boundary traversed by the carpentum and the snake, the boundary on the path between the two suns. Moreover, the bends allow it to take up its flight in the same direction on this side as it began on the other side. It can thus be seen as representing the frontier defended by the Archer Guardian.

2.3 OTHER SYMBOLS. Once we establish the religious nature of these four symbols, other Pictish symbols fall into place with known Celtic religious beliefs. For example, the eagle (or any bird) was regarded as holy because it spoke a language of the otherworld and could also fly to and from the otherworld. Moreover, as noted above, the eagle's vast wing-span represented the sky itself. Other symbols will be treated further below in section 3.2.

Thus, we can readily surmise that the context of these symbols is intended to be religious. Moreover, it is consistent with the general religious beliefs for which we have evidence from throughout the Celtic world (as we would expect—compare

Laing & Laing 1993:21). With the context of situation thus defined, we can now turn our attention to the issue of the grammar of the symbols.

3. THE GRAMMAR OF THE SYMBOLS. As noted above, the symbols are usually written together in pairs, although single symbols and three or even four symbols may also be found (beyond this, we are probably working with groups of two or three). Accordingly, the level of grammar should allow for one unit in isolation, for two units as the preferred configuration, and for an extension to three or four units. Furthermore, where the V-rod is used, the unit must be consistent with the invocation of the Archer Guardian (or of God in the syncretic period) on behalf of a deceased—or in some other dedicatory context, as this interpretation need not be restricted to funerary situations (no more than a cross on a necklace must correspond to one over a grave).

3.1 THE CLAUSE AS ASYMBOLIC UNIT. Quite important to the level of grammar, the unit must also be appropriate to a religious context; and where the V-rod is used, it must be appropriate to what one may be expected to say in an invocation to the Archer Guardian, especially (though not necessarily) where a death has occurred. The only Pagan Celtic invocation extant occurs at the beginning of the Tablet of Chamalières:

> I beseech the very divine, the divine Maponos Avernatis by means of the magic tablet: quicken(?) us, i.e. those (named below) [sic] by the magic of the underworld spirits(?). (Koch 1995:3)

Of course, in the process of syncretism, Paganism adopted rituals and concepts from Christianity 'to supply its own deficiencies' (De Reu 1998:27), especially in matters of liturgy. The closest type of prayer in the Christian tradition to an invocation of the Archer Guardian would be the collect. We find just such a prayer in the introductory collect of the Lorrha (or Stowe) Missal:

> O God, Who to Blessed Peter Thine Apostle didst bestow by the keys of the Heavenly Kingdom, the power to bind and loosen souls, and didst give the office of High Priest, receive our prayers of propitiation and his intercession. We ask O Lord for help that we may be freed from the bonds of our sins through our Lord Jesus Christ Who reigneth with Thee and the Holy Spirit, God throughout all ages of ages. (Dowling 1995)

When we reduce the collect to its essential formula, we find that it always expresses the construction:

> (may) god [who…] do this

Thus, we have two clauses, a petitionary clause with an embedded relative clause modifying the subject being invoked at the beginning of the prayer. The relative clause identifies the attributes of God most appropriate to the type of petition. This opening complex sentence may then be followed by further, related petitions, creating in effect a compound complex sentence.

Such a formula is furthermore quite consistent with Pagan thinking, for we are naming God and asking God to fulfill our petition. The Pagans 'believed in the gods because of the help they offered: they delivered prosperity, strength, a plentiful harvest, good fishing and victory in battle' (De Reu 1998:13–14). Indeed, the Pagan could make a petition directly to the intended divine being, without the awkward requirement of praying only to God while invoking the intercession of a saint.

What is of greatest importance to the linguist, of course, is the fact that we can identify the context of situation as originally a Pagan prayer invoking the Archer Guardian and making a petition. From the internal linguistic, grammatical perspective, moreover, we can further identify the formula as consisting of clauses in a complex or compound complex sentence. Thus, our level of linguistic analysis for the Pictish symbol stones is the clause structure.

This would seem reasonable, for the clause can stand on its own, providing independent information in the appropriate context—an invocation without petition or a petition without invocation. The usual form for a collect, however, is the main clause with the petition and the relative clause with the attributes—thus accounting for the majority of Pictish representations with two symbols/clauses. Where only two symbols occur, they are probably in the form of a general petition, asking for a blessing or protection in accordance with the first sentence of the Lorrha Missal's collect above.

Moreover, in the collect from the Lorrha Missal, we go beyond the general petition, introducing specific petitions appealing to the attributes described in the relative clause. If the Pictish symbol stones do indeed represent Pagan prayers parallel to such collects, this extension of the petition from general to specific would account for the additional symbols. In these cases then, we end up with one or more further symbols/clauses.

3.2. INTERPRETING THE SYMBOL STONES. Thus far, we have the religious context of situation, transparent meanings for the symbols directly connected with the Cult of the Archer Guardian, fairly transparent meanings for some of the other symbols, and the level of grammar consisting of clauses used in invocation and petition.

At this point, then, we can suggest meanings for some of the more transparent combinations. Since these combinations appear on stones often with a funerary intent, we can supply a tentative structure, following that of the collect above, such as the following:

> may the archer guardian, who…,
> protect the departed upon their journey

Thus, the relative clause appropriate for the Z-rod and double-disk may be 'who guards the border between this world and the otherworld,' that for the arch may be

Double-disk and Z-rod	who guards the border between this world and the otherworld
Arch	who guards the entrance to the otherworld
Notched rectangle	who guards the vehicle of translation
Notched rectangle and Z-rod	who guards the vehicle of translation in its journey to the otherworld
Crescent	who watches over the dead/over archers
Pictish Beast	?
Cauldron	who enables resurrection in the otherworld
Snake and Z-rod	who guards the passage to the otherworld
Rectangle	?
Double-disk	who guards the passage between this world and the otherworld
Circle	who ensures life everlasting

Table 2. *Possible interpretations of comginations with the crescent and V-rod.*

'who guards the entrance to the otherworld,' and so forth. Some possible (and highly tentative) interpretations based upon the broader Celtic symbolism are provided in Table 2.

Where the crescent and V-rod come second, we would have to assume (following the formula for the collect) that the prayer is being made to the first element. For example, the combination of the eagle with the crescent and Z-rod may be interpreted as an invocation of the eagle, whose wingspan covers the earth and encompasses the Archer Guardian. Of course, some petitions do not involve the Archer Guardian at all.

4. CONCLUSION. Without a context of situation, the identification of the appropriate level of linguistic analysis for the Pictish symbol stones is quite frankly impossible. While previous researchers may have been able to rule out certain grammatical levels, they have had no way of determining the level itself. Indeed, if one had suggested previous to the identification of the Cult of the Archer Guardian that the interpretation be religious and at the clause level, one could have supported that conclusion with no greater facility than any other.

As modern linguists, we tend to take the context of situation for granted, for our areas of inquiry have become rather limited to oral and written communication with precisely known or 'neutral' context. It is when we go beyond the elicitation of familiar and comprehended verbal language that we see how important knowing the exact context, the exact intent and the exact style of communication really is.

REFERENCES

Allen, J. Romilly & Joseph Anderson. 1903. The early Christian monuments of Scotland. Society of Antiquaries of Scotland. [Rpt. 1993. Balgavies, Angus: Pinkfoot Press.]

Chalmers, Patrick. 1848. *The ancient sculptured stones of the County of Angus*. Aberdeen: Bannatyne Club.

Cummins, W. A. 1995. *The age of the Picts*. Phoenix Mill, Gloucestershire: Sutton.

———. 1999. *The Picts and their symbols*. Phoenix Mill, Gloucestershire: Sutton.

De Reu, Martine. 1998. The missionaries: The first contact between Paganism and Christianity. In *The Pagan Middle Ages*, ed. by Ludo J.R. Milis, trans. by Tarris Guest, 13–37. Woodbridge, Suffolk: Boydell Press.

Driscoll, Stephen T. 1986. Symbol stones and Pictish ethnography: Review of Symbol stones of Scotland. *Scottish archaeological review* 4/1:59–64.

Dowling, Maelruain Kristopher (ed. and trans.). 1995. Celtic missal: The liturgy and diverse services from the Lorrha ('Stowe') missal used by Churches of Ireland, Scotland, Britain, France, Germany, Switzerland, and Northern Italy. http://www.geocities.com/Athens/3374/ stowe.html.

Firth, J.R. 1968. A synopsis of linguistic theory, 1930–55. In *Selected papers of J.R. Firth, 1952–59*, ed. by F.R. Palmer, 168–205. Bloomington: Indiana University Press.

Forsyth, Kartherine. 1996. *The ogham inscriptions of Scotland: An edited corpus*. Ph.D. diss. Cambridge, ma: Harvard University.

———. 1997a. *Language in Pictland: The case against 'non-Indo-European Pictish' (Studia Hameliana 2)*. Utrecht: de Keltische Draak.

———. 1997b. Some thoughts on Pictish symbols as a formal writing system. In *The worm, the germ, and the thorn: Pictish and related studies presented to Isabel Henderson*, ed. by David Henry, 85–98. Balgravies, Angus: Pinkfoot Press.

Green, Miranda J. 1986. *The gods of the Celts*. Godalming, Surrey: Bramley Books.

———. 1995. The gods and the supernatural. In *The Celtic world*, ed. by Miranda J. Green, 465–88. London: Routledge.

Griffen, Toby D. 2000. The Pictish art of the Archer Guardian. Paper presented to the Celtic Studies Association of North America, March 2000. St Louis. [Available at http://www.dubricius.net.]

Hardinge, Leslie. 1972. *The Celtic church in Britain*. Brushton, New York: Teach Services.

Henderson, Isabel. 1971. The meaning of the Pictish symbol stones. In *The dark ages in the Highlands*, ed. by Edward Meldrum, 53–67. Inverness: Inverness Field Club.

Jackson, Anthony. 1984. *The symbol stones of Scotland: A social anthropological resolution of the problem of the Picts*. Stromness: Orkney Press.

Koch, John T. with John Carey. 1995. T*he Celtic heroic age: Literary sources for ancient Celtic Europe and early Ireland and Wales*. Malden, Massachusetts: Celtic Studies Press.

Laing, Lloyd & Jennifer Laing. 1984. Archaeological notes on some Scottish early Christian sculptures. *Proceedings of the Society of Antiquaries of Scotland* 114:277–87.

——— & ———. 1993. *The Picts and the Scots*. Phoenix Mill, Gloucestershire: Sutton Publishing.

Maier, Bernhard. 1997. *Dictionary of Celtic religion and culture*, trans. by Cecil Edwards. Woodbridge, Surrey: Boydell Press.

Meyvaert, Paul. 1992. A new perspective on the Ruthwell Cross: Ecclesia and vita monastica. In *The Ruthwell Cross: Papers from the colloquium sponsored by the Index of Christian art, Princeton University, 8 December 1989*. ed. by Brendan Cassidy. Princeton, New Jersey: Department of Art and Archaeology, Princeton University.

Nicolaisen, W. F. H. 1972. P-Celtic place-names in Scotland. *Studia Celtica* 7:1–11.

———. 1996. *The Picts and their place names*. Rosemarkie: Groam House Museum.

Ritchie, J. N. Graham. 1997. Recording early Christian monuments in Scotland. In *The worm, the germ, and the thorn: Pictish and related studies presented to Isabel Henderson*, ed. by David Henry, 119–28. Balgavies, Angus: Pinkfoot Press.

——— & Anna Ritchie. 1981. *Scotland: Archaeology and early history*. London: Thames and Hudson.

Ross, Anne. 1996. *Pagan Celtic Britain*. Chicago: Academy Chicago Publishers.

Samson, Ross. 1992. The reinterpretation of the Pictish symbol stones. *Journal of the British Archaeological Association* 145:29–65.

Sutherland, Elizabeth. 1994. *In search of the Picts: A Celtic dark age nation*. London: Constable.

———. 1997. A guide to the Pictish stones. Edinburgh: Birlinn.

Thomas, Charles. 1963. The interpretation of the Pictish symbols. *The archaeological journal* 120:31–97.

* * *

EPISODE TRANSITIONS AND THEIR LINGUISTIC ENCODINGS

Shaojun Ji
Department of Linguistics, University of Alberta

INTRODUCTION. A major issue in the study of discourse is how narrative is structured. Numerous studies have already provided good analyses of various aspects of its structure. Longacre (1983:3), for instance, pointed out two main features of narrative discourse. One is what he characterized as the 'contingent temporal succession' of the described actions or events in a narrative. This characterization means that these actions or events are usually not only temporally ordered but also mutually dependent. Such dependency may be created by causality or it may result from a more general sense of coherence. The other main feature of narrative discourse, according to Longacre, is its 'agent orientation'. This orientation reflects the fact that narratives are typically built around human agents or other animate beings.

It has also been established that narrative is not composed merely of sequences of sentences, but is hierarchically organized into intermediate units such as episodes or paragraphs (Bamberg & Marchman 1991; Black & Bower 1979; van Dijk 1982; Grimes 1975; Hinds 1977, 1979; Longacre 1979; Pu, Prideaux & Stanford 1992; Tomlin 1987). This fact is generally in line with results from studies that are more focused on the issues of how thematic continuity is maintained and how thematic discontinuities are signaled in the narrative type of discourse (Brown & Yule 1983; van Dijk & Kintsch 1983; Givón 1983, 1984; Virtanen 1992).

1. TOWARDS AN EXPERIMENTAL STUDY. Based on all the research mentioned above, the present study carried out an empirical investigation of language users' conception of episode transition and its role in the organization of narrative discourse. More specifically, the present study seeks to determine what the important elements are that form the basis of our conception of episode transitions and how such transitions are linguistically expressed.

Empirical support for the validity of episode has been provided by psycholinguistic studies which indicate that people recalling stories treat information of an episode as an integral unit (cf. Black & Bower 1979; Glenn 1978; Mandler 1978; Thorndyke 1977), a result which is further supported by studies that have found the episode-shift effect, i.e., readers pause longer in processing episode-initial sentences that introduce an episode shift than episode-internal sentences (Haberlandt, Berian & Sandson 1980; Mandler & Goodman 1982). Although an episode may not have its boundary marked as the indented paragraph, these studies suggest that the boundary of an episode is recognizable on thematic grounds.

Equally important contributions to our understanding of the structure of narrative have been made by studies using other methodologies. Based on his study of *The Pear Story* data, Chafe (1979, 1980) not only demonstrated the relationship between major thematic breaks and processing difficulty by using hesitation as evidence, he also pointed out that such thematic breaks tend to be associated with a change in time, place, character configuration, or event structure. Because the focus of his study was to show the relationship between the difficulty of oral production and points of thematic shifts, he did not discuss in a comprehensive way the forms of linguistic expressions that are used at these points of shift.

Focusing their attention on the discourse function of preposed subordinate clause of time in their experimental research, Prideaux & Hogan (1993) found that in both oral and written narratives, the preposed subordinate clause of time (a marked structure) occurred statistically far more frequently at the beginning of an episode than the nonpreposed one (an unmarked structure). The nonpreposed subordinate clause of time tended to occur far more frequently within an episode. Their analysis indicates that because the preposed subordinate clause is a good device for thematic reorientation, it is often used to code the beginning of a new discourse unit. Givón (1993:315) made the same observation as Prideaux and Hogan (1993) and, in addition he also suggested that preposed participial phrases and preposed prepositional phrases may serve the same discourse function.

2. IDENTIFYING EPISODE TRANSITIONS: A SEGMENTATION TASK. As the survey indicated above, the previous studies have, for various reasons, suggested a correlation between episode transitions and major changes of such discourse elements as time, place, or participant. However, a study is needed to show the empirical situation of this correlation based on episode boundaries as identified by ordinary language users. If we can acquire some empirical evidence about ordinary language users' view of how narrative develops from one episode to another, we will better understand how narrative is structured. The following is the hypothesis for the present study:

> In narrative discourse, episode transitions are primarily defined by major changes in time, place, or participant. When the contents of a narrative lack such changes as anchorable points of transition, the episode boundary will become less clear.

To test this hypothesis, an experiment was conducted in which 20 native speakers of English and 20 native speakers of Chinese were asked to segment a few narrative texts in their own languages. The narrative texts used for both language groups were personal accounts of past experience. The three English narrative texts used as stimuli were all from *The Reader's Digest*. The four Chinese narrative texts were from two major literary magazines published in China. Each text was given to the participants in the experiment in an unparagraphed form but with a title. The participants were asked to divide the texts into episodes by placing a vertical stroke before the first word of the episode.

An episode was defined to the participants as a portion of a narrative that relates to an event or a series of connected events and forms a coherent unit in itself.

3. RESULTS AND ANALYSES.

3.1. SENTENCES IN THE ENGLISH HIGH-RATE GROUP. After the segmented narratives were collected, all the identified sentences were tabulated and those identified by 70% or more of the participants were picked out and analyzed as representative episode-initial sentences because 70% stands for a clear majority. These sentences are categorized as the *high-rate group*. Excerpts 1, 2, and 3 provide some examples from this group:

(1) a. ***Now, hearing of her elephantine raspberries,*** I went immediately to Ellie's quarters. She looked at me with moist, gray eyes and flapped her ears. "Praaa," went the trunk. (followed by the narrator's temporary diagnosis and decision)
  b. ***Back at my office,*** I phoned a surgical-instrument manufacturer. I asked to borrow a flexible endoscope. (followed by an explanation of the use of the instrument and the result of getting it)
  c. ***The next morning*** I injected Ellie with anesthetic. Five minutes later she sank to her knees, unconscious, her trunk still blaring when she exhaled. (followed by the narrator's examination of the inside of Ellie's trunk)

(2) a. ***Before attempting to scale the heights in the real world***, I had gone to Chelsea Piers, a sports complex in Manhattan, to train on their climbing wall. For fifty bucks, I spent two hours learning to tie the classic climbing knots and use the equipment. (followed by an account of how the narrator became motivated in climbing a real mountain)
  b. ***Forty feet up a novice climb called Jackie***, I got stuck between a crack and an overhang. I was sure that if I moved either of my feet or my paws, gravity would have its way with me. (followed by a description of the narrator's difficult situation and her further actions of climbing)
  c. ***After I conquered Jackie***, Paul took me over to Three Pines, an easier multiple-pitch climb—which means climbing cliffs that are higher than your rope is long. (followed by an account of the climbing process)

(3) a. ***On Monday morning*** Ellie came to see me again. And again on Tuesday, Wednesday, Thursday and Friday. Each time I looked at the trunk and told her how well she was doing. On the morning of her tenth office visit, I snipped out the steel sutures. The wound had healed and would leave hardly any scar. Ellie must have decided that the removal of the stitches indicated that I was signing her off, for she didn't come the next day, nor ever again.
  b. ***A fellow vet once told me that his father had been prescribed champagne after an operation for nasal polyps.*** "It sounds like sensible medicine," he had remarked. I agreed. So a few days later, I bought a bottle to split in celebration with Ellie's keeper. The patient, however, did not go unrewarded. Ellie got a large, iced currant bun.

According to the hypothesis of this study, if episode transitions are characterized by major changes in time, place, or participant, the episode-initial sentences should contain information about these changes. So, an important part of the analysis of these sentences is to examine them one by one and find out if they are characterized by such changes. Within each episode-initial sentence the sentence-initial structures are the focus of our attention. From the point of view of textual development, the initial position of the transitional sentence is crucial in the sense that it is easy for the element placed here to connect the sentence with what has preceded in the previous thematic unit. It is also a natural starting point for the rest of the sentence or the whole unit it leads.

Table 1 summarizes the correspondences between sentence-initial structures in the English high-rate group and the three types of discontinuities:

In tabulating the changes in time, place, or participant, we set up three functional categories, namely temporal, spatial, and participant discontinuities. There are altogether 14 sentences in the high-rate group. Table 1 records the correspondences between the sentence-initial structures and certain discontinuities in 13 sentences. Because one of them has two initial structures, Table 1 actually contains 14 tokens of structure.

A case of temporal or spatial discontinuity corresponds to a sentence-initial structure indicating a clear change in time or place. For instance, *now* in (1)a is recorded as adverb expressing a temporal discontinuity, whereas *back at my office* in (1)b is treated as a case of prepositional phrase indicating a spatial discontinuity. *The next morning* in (1)c is regarded as a noun phrase informing a temporal discontinuity. If a sentence starts with a new participant in its subject position, it is treated as a case of participant discontinuity. *A fellow vet* in (3)b is an instance of such a case.

A case of temporal discontinuity may also be a sentence-initial participial phrase, a prepositional phrase, or a subordinate clause such as *hearing of her elephantine raspberries* in (1)a, *before attempting to scale the heights in the real world* in (2)a, and *after I conquered Jackie* in (2)c. What these structures have in common is the indication of a particular time by referring to a specific event or activity in the previous episode. According to Givón (1993:315), all these sentence-initial structures have their coherence links in two directions, anaphoric and cataphoric. Their anaphoric links can reach back to thematic information anywhere in the preceding episode. Their cataphoric links anchor themselves nicely to the main clause, which then launches the new episode. (One transitional sentence in the English high-rate group had to be analyzed on its own.)

As recorded in Table 1, the initial structures in all the English high-rate group sentences but one are found to carry the information indicating major changes in time, place, or participant. A very prominent feature about the structures expressing these changes is that except in the case of participant discontinuity in Table 1, which is expressed by the indefinite noun phrase in the subject position (***A fellow*** *vet once told me…*), all the other structures informing changes in time or place are preposed and therefore 'marked' structures. These preposed structures are regarded as marked because the canonical positions for these structures in English sentences are at the

Syntactic Categories	Functional Categories		
	Temporal Discontinuity	Spatial Discontinuity	Participant Discontinuity
Adverb	2		
Prepositional Phrase	3	2	
Participial Phrase	1		
Noun Phrase	3	1	1
Subordinate Clause	1		
Total	10	3	1

Table 1. *Sentence-initial structures expressing different types of discontinuities (based on sentences of the English high-rate group: 13/14=93%).*

end of the sentence (e.g., Clark & Clark 1977; Givón 1983; Quirk et al. 1985). The consistent use of these marked structures constitutes a very prominent feature about the way temporal and spatial discontinuities are expressed in the episode-initial sentences. This observation supports Prideaux & Hogan's (1993) finding about the frequent use of the marked subordinate clause of time at episode-initial position. It is also compatible with Virtanen's (1992) data of clause-initial adverbials used at the beginning of textual units. Following Givón (1993), Prideaux & Hogan (1993), I believe that the consistent use of marked structures at episode-initial position is functionally motivated. This functional motivation is to signal the start of a new episode by referring to the changes in time and place.

3.2. SENTENCES IN THE CHINESE HIGH-RATE GROUP. The segmented Chinese texts obtained from Experiment 1 were processed in the same way as their English counterparts. The episode-initial sentences in the Chinese high-rate group may be exemplified by those occurring in the two excerpts below:

(4) a. ***shengcheng xia lai de nei ge zhiqing*** *kaishi*
provincial capital descend come NOM[1] **that** CL **educated youth** initially
ting bu dong houlai mingbai le juede youqu jiu yixiao.
listen not understand afterwards understand PFV feel funny just laugh
'**The educated youth that came from the provincial capital** did not understand (the term) in the first place. Understanding it later he thought it was funny and laughed about it.' (followed by an account of how the young man differed from the villagers in speech and for that reason was laughed at by them)

b. ***ruogan nian hou,*** *shengcheng xia lai de nei ge zhiqing*
several years later provincial capital descend come NOM that CL educated youth
bujin tingdong le xiangcun de fangyan liyu, erqie po you yanjiu.
not only understand PFV countryside ASSOC dialect slang but also quite have research
'**Several years later,** the educated youth that came from the provincial capital not only understood the dialect and slang in the rural area but actually knew quite a lot about them.' (followed by an account of how he got the nickname *Wutian* from the villagers)

(4) c. ***Wutian*** *qing shou, baijing, tianting baoman, shouzhi shuicong yiyang.*
Wutian pretty thin fair and clean forehead plump finger water onion same shape
'**Wutian** is pretty thin and looks fair and clean. His forehead is plump. His fingers are just like water onion.' (followed by an account of how the local people accepted him)

d. ***zhiqing gang xia lai de shihou,*** *cunliren you xinxian gan,*
educated youth just descend come NOM time villager have fresh interest
you shi mei shi dou xihuan zuan zhiqing wu zuozuo,
have thing not have thing all like go into educated youth rooms sit a while
'**When the educated youths first came**, the villagers had great interest. Whether or not they had anything to do (with the educated youths), they all liked to go into their rooms and sit for a while.' (followed by an account of how the young man *Wutian* had his first romantic experience with a local girl)

(5) a. ***dao Changdu hou,*** *wo qing xian renshi ju zhang*
reach Changdu after I ask county personnel bureau director
Xiangba zuo fanyi, ta xinran daying.
Xiangba to be interpreter he gladly agree
'**After (we) arrived at Changdu**, I asked the director of the personnel bureau of the county to be my interpreter. He gladly agreed.' (followed by a description of the condition of the road in Tibet)

b. ***chu Changdu bu yuan,*** *jiu likai le Chuangzang gonglu de zhugandao.*
leave Changdu not far soon leave PFV Sichuan-Tibet road NOM main line
'**Soon after we were out of Changdu**, (we) left the main line of the Sichuan-Tibet road.' (followed by a description of the condition of the road ahead)

c. ***bangwan shifen,*** *women dao le Caiwei qu.*
evening time we reach PFV Caiwei district
'**In the evening**, we arrived at Caiwei district.' (followed by a description of the place where the district administration is located)

d. ***quli zhi you jige ren zai.*** *quli*
district administration only have a few people remain district administration
de lingdao dao xianli kai hui qu le.
ASSOC officials to county hold meeting go PFV
'**There are only a few people left at the district administration.** The officials of the district administration have gone to the county to attend a meeting.' (followed by an account of their experience of having dinner at the local government)

e. ***chi le wanfan,*** *Gesang zai yuanzi li ca che,*
eat PFV supper Gesang at courtyard in clean car
wo he Xiangba sanbu zou chu.
I with Xiangba take a walk go out
'**After we had supper**, Gesang was cleaning his car in the courtyard. I and Xiangba went out in the manner of taking a walk.' (followed by an account of visiting a local Tibetan in his house)

In the analysis of the Chinese high-rate group sentences, (4)a represents a case where the subject noun phrase (modified by a relative clause) introduces a new

Syntactic Categories	Functional Categories		
	Temporal Discontinuity	Spatial Discontinuity	Participant Discontinuity
Adverb	3		
Prepositional Phrase	5		4
Participial Phrase			5
Noun Phrase	3	2	
Subordinate Clause			2
Total	10	2	11

Table 2. *Sentence-initial structures expressing different types of discontinuities (based on sentences in the Chinese high-rate group: 22/23=96%).*

participant into the story. (4)c is a case of subject noun phrase serving the same function but has no modification. In cases where the subject noun phrases are modified, either by a relative clause or a genitive phrase, the modifying structures often provide the necessary background information or convenient reference point for the introduction of the new participant. (5)d is an existential sentence. An existential sentence in Chinese such as (5)d starts with a location, which is conceptualized as a possessor. What exists at this location is presented as the possessed. Despite this, we do not tabulate this initial noun phrase of location as a structure involving a change of location because in these circumstances the location has already been activated in the previous episode. Instead we treat the whole sentence pattern as a special form for introducing new participants, which, unlike the subject noun phrases in normal sentences, comes at the end of the sentence. Following Lambrecht (1988:15–16), I regard the sentence-initial noun phrase of location as serving the function of anchoring the new referent in the discourse.

Temporal discontinuity is also expressed in various ways. (4)b and (5)c are instances of noun phrases indicating major changes in time. As a subordinate clause, (4)d and (5)e each provides specific information of time relating to a specific event. (5)a and b are both subordinate clauses of time in form, though (5)b contains no temporal conjunction. However, the changes in place indicated by the verb phrases in both clauses seem to be more informative than the times in relation to them. Based on this reasoning, (5)a and b are categorized as two cases of subordinate clauses responsible for spatial discontinuities.

Table 2 records the correspondences between initial structures of the 22 sentences in the Chinese high-rate group and the specific discontinuities they indicate. The Chinese high-rate group consists of 23 sentences. Among the 22 sentences recorded in Table 2, one sentence contains two initial structures—as a result, 23 tokens are listed. One sentence in the Chinese high-rate group is categorized as a case of thematic discontinuity. Because it is the only case of thematic discontinuity it is treated on its own.

From the above discussion we can see that most of the sentences in the Chinese high-rate group are found to be characterized by major changes in time, place, or participant. Except in the two existential sentences, the change in participant is consistently expressed by the subject noun phrase at sentence-initial position. Changes in time and place, on the other hand, are expressed by such sentence-initial structures as adverbs, noun phrases, and subordinate clauses. The expressions of these changes in the Chinese data are quite compatible with those observed in the English data in the sense that as carriers of information about important changes in time, place, or participant they need to be placed right at the very beginning of an episode (unless a special sentence pattern such as existential sentence or cleft sentence is used).

In the analysis of the English data, the sentence-initial structures responsible for indicating changes in time and place are characterized as marked (in comparison with their canonical positions at the end of the sentence). In the case of Chinese, however, the same structures regularly occur at sentence-initial position as 'topics'(Chafe 1976: 50–51; Li & Thompson 1981:94–95) since Chinese is a topic-prominent language (Li & Thompson 1981). The functional constraint that the information about important changes in time and place is crucial for the introduction of a new episode provides a good perspective for similar syntactic behaviors from two syntactically very different languages.

3.3. COMPARING HIGH-RATE GROUPS WITH LOW-RATE GROUPS. Although my investigation was focused on sentences in the high-rate groups I also had a close look at sentences identified with 40%–65% agreement rates in the low-rate groups of the two languages. What I find is that sentences in the low-rate groups are also characterized by the three types of discontinuities. But they are indicators of local changes within the contexts of episodes set by the transitional sentences of the high-rate groups. Another finding is that many sentences in the low-rate groups may be characterized as introducers of new themes. Considering this fact and the fact that only one case of thematic discontinuity was identified in the high-rate groups, sections specifying certain themes seem more likely to be regarded as embedded units within episodes.

4. SUMMARY. As the above analysis shows, the result of the segmentation task supports the hypothesis of this study. In more specific terms, the result may be summarized as follows:

First, in identifying episode transitions both English- and Chinese-speaking subjects used the same strategy: they looked for sentences that involve changes in time, place, or participant. This strategy indicates that these changes are natural indicators of episode transitions in narrative texts and as such they are perceived by language users.

Second, episode transitions are primarily defined by major changes in time, place, or participant. The low-rate cases typically involve local changes in these dimensions. The lesser agreement in identifying the low-rate cases implies that they represent less clear cases of episode transitions.

Third, in episode-initial sentences the changes in time and place are typically encoded in preposed structures whereas the changes in participant are often indicated by subject noun phrases.

Finally, the result of the experiment has made us aware that if changes in time, place, or participant generally correlate with episode transitions, it means that the discourse elements of time, place, or participant are also important to our conception of episode as a coherent unit. In other words, our sense of episode as a thematically coherent unit is derived from our evaluation of the roles played by these elements in the formation of the mental representation of an episode. This evaluation is cognitively based in the sense that we (together with other people) engage in many different activities in different periods of time at different places every day and our common view as to which activities should belong to one rather than another 'episode' of our daily life seems to rely heavily on the parameters of the time or place of the activities as well as who participate in them.

---

[1] Abbreviations: ASSOC – associative; CL – classifier; NOM – nominalizer; PFV – perfective aspect

REFERENCES

BAMBERG, MICHAEL & VIRGINIA MARCHMAN. 1991. Binding and unfolding: Towards the linguistic construction of narrative discourse. *Discourse processes* 14: 277–305.

BLACK, JOHN G. & GORDON H. BOWER. 1979. Episodes as chunks in narrative memory. *Journal of verbal learning and verbal behavior* 18: 309–31.

BROWN, GILLIAN & GEORGE YULE. 1983. *Discourse analysis*. Cambridge: Cambridge University Press.

CHAFE, WALLACE. 1976. Givenness, contrastiveness, definiteness, subjects, topics and points of view. In *Subject and topic*, ed. by Charles Li, 25–55. London: Academic Press.

———. 1979. The flow of thought and the flow of language. In *Discourse and syntax*, ed. by Talmy Givón, 159–81. New York: Academic Press.

———. 1980. The deployment of consciousness in the production of a narrative. In *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, ed. by Wallace Chafe, 9–49. Norwood NJ: Ablex.

CLARK, HERBERT H. & EVE V. CLARK. 1977. *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace, Jovanovich.

VAN DIJK, TEUN A. 1982. Episodes as units of discourse analysis. Deborah Tannen, Ed., Analyzing discourse: text and talk, 177-95. Georgetown University Round Table on Languages and Linguistics 1981. Washington DC: Georgetown University Press.

——— & WALTER KINTSCH. 1983. *Strategies of discourse comprehension*. New York: Academic Press.

GIVÓN, TALMY. 1983. Topic continuity in spoken discourse. In *Topic continuity in discourse: A quantitative cross-language study*, ed. by Talmy Givón, 343–63. Amsterdam: John Benjamins.

———. 1984. *Syntax: A functional-typological introduction*, vol. 1. Amsterdam: John Benjamins.

———. 1993. *English grammar A function-based introduction*, vol. 2. Amsterdam: John Benjamins.

Glenn, Christine G. 1978. The role of episodic structure and of story length in children's recall of simple stories. *Journal of verbal learning and verbal behavior* 17:229–47.

Grimes, Joseph E. 1975. *The thread of discourse*. The Hague: Mouton.

Haberlandt, Karl, Clair Berian & Jennifer Sandson. 1980. The episode schema in story processing. *Journal of verbal learning and verbal behavior* 19:635–51.

Hinds, John. 1977. Paragraph structure and pronominalization. *Papers in linguistics* 10: 77–99.

———. 1979. Organizational patterns in discourse. Talmy Givón, Ed., Discourse and syntax, 135–56. Syntax and semantics 12. New York: Academic Press.

Lambrecht, Knud. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*. Berkley: Berkley Linguistics Society.

Li, Charles & Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California.

Longacre, Robert E. 1979. The paragraph as a grammatical unit. In *Discourse and syntax*, ed. by Talmy Givón, 115–34. New York: Academic Press.

———. 1983. *The grammar of discourse*. New York: Plenum Press.

Mandler, Jean M. 1978. A code in the node: The use of a story schema in retrieval. *Discourse processes* 1:14–35.

——— & Marsha S. Goodman. 1982. On the psychological validity of story structure. Journal of verbal learning and verbal behavior 21:507–23.

Prideaux, Gary D. 1994. Two complementary approaches to discourse analysis. *Linguistica atlantica* 16: 113–33.

——— & John T. Hogan. 1993. Markedness as a discourse management device: The role of alternative adverbial clause orders. *Word* 44(3):397–411.

Pu, Ming-Ming, Gary D. Prideaux & Lois M. Stanford. 1992. Episodes as memory units in discourse representation: Anaphor use in English and Mandarin. *Linguistica atlantica* 14:89–102.

Quirk, Randolf, Greenbaum Sidney, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Thorndyke, P. W. 1977. Cognitive structure in comprehension and memory of narrative discourse. *Cognitive psychology* 9:77–110.

Tomlin, Russell S. 1987. Linguistic reflections of cognitive events. In *Coherence and grounding in discourse*, ed. by Russell S. Tomlin, 455–79. Amsterdam: John Benjamins.

Virtanen, Tuija. 1992. *Discourse functions of adverbial placement in English: Clause-initial adverbials of time and place in narratives and procedural place descriptions.* Åbo/Turku, Finland: Åbo Akademi University Press.

* * *

POETIC LICENSE, EMBELLISHED BY MUSIC
(in memory of Robert Di Pietro, president of LACUS in 1990)

Saul Levin
State University of New York at Binghamton

DEPARTING FROM EVERYDAY USAGE is a nearly universal phenomenon for poets, whether we call it poetic style or poetic language. This can go much further when the words are set to music, as I will show in examples from well-loved masterpieces.

1. THE FIRST ONE is by Gilbert and Sullivan (1878), Sir Joseph Porter's song in *H.M.S. Pinafore*:

> When I was a lad I served a term
> As office boy to an attorney's firm.
> I cleaned the windows and I swept the floor,
> And I polished up the handle of the big front door.
> (CHORUS: He polished up the handle of the big front door.)
> I polished up the handle so care-ful-lee
> That now I am the ruler of the Queen's Na-vee.
> (CHORUS: He polished up the handle to carefullee
> That now he is the ruler of the Queen's Navee.)

In this 'patter' song the vocabulary is plain, unpoetic, and the meter purposely uneven: artfulness is therefore displayed all the more in the rhyming of the couplets. Then, in the last couplet, the world *ruler*—though a little on the literary side—is a little less fancy a title than *First Lord of the Admiralty* would be with its archaic sound, and *the Queen's Navy* is an up-to-date substitute for *the Admiralty*. But only a comedian would rhyme 'Navee' with 'carefullee'; the unstressed *-y* in a two-syllable word cannot otherwise be treated as a rhyming vowel[1].

However, the rhyme *carefullee : Navee* is quite tame, compared to the one in the third stanza (and thereafter)[2]:

> In serving writs I mades such a name
> That an articled clerk I soon became;
> I wore clean collars and a bran' new suit
> For the pass examination at the Institute
> (CHORUS: For the pass examination at the Institute.)
> That pass examination did so well for me
> That now I am the ruler of the Queen's Na-vee.

(CHORUS: That pass examination did so well for he [!]
That now he is the ruler of the Queen's Navee.)

Nothing in the English language would bring on 'so well for he', except a choral refrain echoing 'so well for me'. Truly a trick of genius, however outrageous.

The idea of taking such liberties occurred to Gilbert and Sullivan perhaps from older but lesser versifiers in the English theater, whose ephemeral compositions were soon forgotten. But another influence came through Gilbert's own labor, when he dipped into the libretti of Italian operas. For there was a sizable public in London for opera, both in the original and in translation. Unlike the fictitious Joseph Porter, who grew rich in the practice of law, the real man Gilbert after passing the bar could not earn a living at it; eventually he put his legal knowledge to better use in numerous scenes of his comic operettas[3]. Meanwhile he found some odd jobs by writing an English text based on operas already popular on the Continent. His first minor success in that line was a translation of a French soprano aria from *Manon Lescaut* by Auber[4]. After that, but before he began collaborating with Sullivan, he wrote *Dulcamara* (1866), a take-off on Donizetti's comic masterpiece *L'Elisir d'amore*, which had become a familiar classic[5].

As is well known, Gilbert and Sullivan's operettas often parody the style of Italian, French, or German opera—tragic opera still more than comic. This extends, as I maintain, to details of meter and musical diction. I have not discovered an exact Italian or other precedent for rhymes like 'so well for he / …the Queen's Navee'; but many devices of Italian grand opera did attract these talented Englishmen and are furthermore worth appreciating in their own right, even apart from the use that Gilbert and Sullivan made of them. For from European culture of the nineteenth century nothing has proved more enduring than opera, nor broader, more nearly universal in its human appeal[6].

II. WORDS SUCH AS *CAREFULLY*, stressed on the antepenultimate syllable, are less important in English than in Italian versification, where they are termed *voci sdrucciole* (literally 'slippery words'). For according to the favorite meters of lyric poetry, the opening line of a stanza or strophe ends in a *voce sdrucciola*, and some or all of the odd-numbered lines thereafter end in another *voce sdrucciola*, NON-RHYMING as a rule[7]; e.g., Leonora's cavatina in *Il Trovatore* (Act I, scene 2) by Verdi:

Tacea la notte placida	('Still was the night, placid
e bella in ciel sereno;	and fair in a sky serene;
la luna il viso argenteo	the moon her silvery face
mostrava lieto e pieno…	showed gay and full…')

The two unstressed syllables of *placida*, and again *-teo* in the third line, bring out poetically the continuing suspense, as the stanza is not yet ready to come to a close. The composer too liked, still more, to dwell upon those syllables[8].

Later on, in Leonora's half of the frenzied duet with the villainous Count di Luna (Act IV, scene 1), Verdi gave a HIGHER pitch to the final syllable of each *voce sdrucciola*:

Mira, di acerbe lagrime	('Look, of harsh tears
spargo al piede tuo un rio.	I shed at thy[9] foot a stream.
Non basta il pianto? Svenami,	Weeping is not enough? Stab me,
ti bevi il sangue mio.	drink up my blood.
Calpesta il mio cadavere	Trample my corpse
ma salva il trovator.	but save the troubadour.')

The untoward emphasis upon the last syllable of *lágrimé* 'tears' after the normal stress upon *la-*, adds immensely to the effect[10]. Even more, *svénamí* pours out the heroine's anguish. For normally the suffixed pronoun *-mi* would be weak, as in *baciami* 'kiss me'—whereas to give emphasis, 'kiss ME' (not someone else), calls for *bacia me*, using the independent form of the object pronoun. But *svena me* 'stab ME' would in any case be incompatible with the meter at this point of the stanza; nevertheless, *svenami* communicates most pathetically, as a *voce sdrucciola* but with *-mi* sung on a prolonged F after *svena-* on E flat. Music triumphs over the rules of the Italian language[11]. However, when the word is repeated five times in the rest of the duet, *-mi* resumes its normal phonetic status[12].

The liberty that Verdi took with his national language, by varying the stress upon *svenami*, had one supreme precedent: Mozart's 'Alleluja' in the motet *Exsultate, jubilate*. This, to be sure, is a very special case; for in Church Latin, and before that in the Greek of the Septuagint, the Hebrew expression הללו יה {HLLW YH} had always been treated as untranslatable, even though the meaning is clear enough in Hebrew: the verb 'Praise ye' with its object the shorter form of the divine name (which otherwise is written as the Tetragrammaton יהוה {YHWH})[13]. Moreover, the ancient synagogues whose main language was Greek cherished this one familiar bit of Hebrew from the Psalms. Thereafter the Greek and other Christian churches carried on the chanting of Psalms, and thus *Alleluia* spread in time to every European language. Accenting the vowel [u] in the penult of this long foreign word was congenial to Greek, and likewise to Latin; but it lacks the sort of inherent fixity that belongs to a native word such as *fortūna* or *pictūra*. So nothing blocked Mozart's impulse to move right on from *Allelúja* to *Allélujâ*, and back and forth many more times[14].

Verdi exercised a different sort of license to suggest a rhetorical tone in *La Traviata* (Act II, scene 1), where the elder Germont, the father of Violetta's lover, remonstrates with her to give him up for the sake of his family:

Pura siccome un angelo	('Pure as an angel
Iddio mi diè una figlia.	God gave me a daughter.
Se Alfredo nega riedere	If Alfred will not come back
in seno alla famiglia,	into the bosom of the family,
l'amato e amante giovine	the beloved and loving youth

cui sposa andar dovea	whose bride she was to be
or si ricusa al vincolo	now draws back from the bond
che lieti ne rendea.	that made us happy.
Deh non mutate in triboli	Pray, do not change into thorns
le rose dell’ amor.	the roses of love.
A’ prieghi miei resistere	To withstand my pleas
non voglia il vostro cor.	let not your heart wish.’)

The natural stress upon the first syllable of *pura* sets a pattern for ['iddio] instead of the normal [idd'io]. Similarly *se Al-*, which counts as one syllable metrically, gets the beat that would otherwise come on the middle syllable of *Alfredo*. Then *l’ámato* is handled just like *Iddio*; and *lé rose* also with unusual stress upon the article. But Verdi applies the device flexibly, not to each and every verse; and indeed, when ‘le rose dell’ amor’ is repeated, the natural pronunciation *le rose* reasserts itself[15]. For here the speaker shifts from deliberate argument to a heart-to-heart appeal; the composer enhances it by inserting in the last verse *no, no* before *non voglia*—regardless of the poetic meter—and again at the very end.

The closing scene of *Il Trovatore* shows how the poetic libretto—before the composer took over—played neatly upon two contrasting infinitives, as the poison that Leonora had swallowed seizes her and she says to her truelove Manrico:

Prima che d’altri vivere	(‘Sooner than to live as another’s
io volli tua morir.	I wished to die thine.’)

The morphology of Italian, or of the standard Tuscan dialect, came with *vivere*, a *voce sdrucciola*, and *morire*, a *voce piana* (a ‘plain word’—i.e., stressed on the next-to-last syllable). The vowel at the end of either infinitive may be dropped in a NON-FINAL position. Moreover, the special rules of lyric poetry require at the end of a stanza a *voce tronca* (a ‘clipped word’) with a final accented vowel or else with a liquid or nasal consonant[16]. So the poet Cammarano reinforced the opposite meaning of the infinitives by putting *vivere* at the end of one line and balancing it with the clipped *morir*. Verdi went further, spreading *vivere* over six notes, from C up to A flat, and then limiting *morir* to two notes: *mo-* on the same C but *-rir* down to B flat.

III. WAGNER, WRITING AS A GERMAN CRITIC of the Italian composers, censured them for distorting the native sound of the words so as to fit the beat of the music[17]. This, however, was due in part to a long-established principle of Italian poetry, that the meters should be flexible within limits; e.g., an iambic verse allows the substitution of a trochee in one or two measures, except at the end, and thus the pattern from verse to verse varies subtly and pleasantly. But when combined with a tune, the words become subordinate; if the wrong syllable, here and there, gets the beat, it scarcely hurts our enjoyment of the song. In this cabaletta from *Don Pasquale* by Donizetti:

So anch' io la virtù magica	('I too know the magic power
d'un guardo a tempo e loco;	of a glance at the [right] time and place;
so anch' io come si bruciano i cori a lento foco.	I too know how hearts burn in slow fire.
D'un breve sorrisetto	Of a brief little smile
conosco anch' io l'effetto,	I too know the effect,
di menzognera lagrima,	of a false tear,
d'un subito languor…	of a sudden faint…)

the rollicking tune, in the third line, slides past *come* 'how' as though it were *comé*. In any case the melody suits the mood handsomely; I cannot wish that Donizetti or his librettist had thought of something to obviate this mild irregularity, a trochaic second measure instead of an iamb[18].

Music in Germany was less inclined than in Italy to cause a shift from one syllable to another, because of a basic phonological difference between the languages. Every syllable in Italian has a full vowel, whereas German words of two or more syllables have one with a weak vowel; therefore to stress the second syllable of *komme* or *kommen* '(to) come', as was done in the Italian [kom'e], would be grotesque. Nevertheless, German COMPOUNDS, which have a minor stress upon one syllable of the second component, lend themselves to uneven treatment in song. Schubert, before Wagner's time, took advantage of this in his setting of Goethe's ballad *Der Erlkönig*. Since the poet had varied the compound between three syllables and four (*Erlenkönig*):

Siehst, Vater, du den Erlkönig nicht	('Seest thou not, father, the Erlking?
den Erlenkönig mit Kron und Schweif?	the Erlenking with crown and tail?')[19]

—with or without the definite article:

Mein Vater, mein Vater, und hörest du nicht	('…and hearest thou not
was Erlenkönig mir leise verspricht?	what Erlenking softly promises me?')

the composer was entitled to a further variation, weakening the stress on the first syllable and placing the beat instead right after it:

Mein Vater, mein Vater, und siehst du nicht dort	('…and seest thou not there
Erlkönigs Töchter am düstern Ort?	Erlking's daughters in the gloomy place?')

Metrically the verse with *Erlkönigs* is quite parallel to this sequel:

Mein Vater, mein Vater, jetzt faßt er mich an!	('…now he grabs me!
Erlkönig hat mir ein Leids getan!	Erlking has done me a hurt!')

but here Schubert returns to the normal stress.

With his sense of drama he was free to exploit such possibilities of his language, somewhat as Verdi or Donizetti did in working with an Italian text. Schubert added the power of musical pitch to build toward a climax: The boy's first utterance begins on C, his next couplet on D, the third on E, and the last one on F as his terror keeps mounting. The music prepares us emotionally, more than words alone could, for the closing sentence: 'In seinen Armen das Kind war tot' ('In his arms the child was dead').

IV. MY FINAL EXAMPLE is from a French tragic opera, *Les Pêcheurs de perles* by Bizet, where—in a manner quite different from Gilbert and Sullivan—the rhyme that ought to close the stanza gets shifted so as not to make it the last word:

Je crois entendre encore	('I seem to hear again,
caché sous les palmiers	hidden under the palm trees,
sa voix tendre et sonore	her voice tender and tuneful
comme un chant de ramiers.	like a song of wood doves.
O nuit enchanteresse,	O enchantress night,
divin ravissement!	divine rapture!
O souvenir charmant,	O charming memory,
doux rêve, folle ivresse!	sweet dream, mad inebriation!
Aux clartés des étoiles	By the twinkling of the stars
je crois encor la voir	I seem again to see her
entr'ouvrir ses longs voiles	half-opening her long veils
aux vents tièdes du soir.	to the warm evening breezes.
O nuit enchanteresse,…	O enchantress night,…)

The refrain has folle *ivresse* rhyming with *enchanteresse*; but the score changes the last line to end with *doux rêve*, which does not rhyme with anything but expresses better the mood of the hero Nadir. Although the French noun *ivresse* does not convey the same moral disapproval on the part of the speaker as its literal equivalent drunkenness in English, still the composer was wise to depart from the order of words given by the librettist[20]. For as the lover reminisces, his senses recapture a delicately blended REVERY; the exquisite sexual hints float down on him one by one. Some other composer, responding to the same words, might have scored them to make *folle ivresse* not only the very climax but the very last word; however, the melody that came to Bizet's ears ends more gently[21].

Music and poetry are there to be enjoyed, and appreciated without requiring an analysis of the art that has gone into them. But if our intellectual curiosity is aroused, it can contribute something positive, explaining how the human voice finds the ways to deliver its message beautifully, and thus to surpass the ordinary level of communication through language[22].

[1] The vowel at the end of *carefully* might receive a secondary stress, being distant enough from the stress on the opening syllable, and would thus make an acceptable rhyme with [-'i:] even in serious poetry.—*Nauie* or *nauye* in Middle English kept the stress on [i] as in Old French; Chaucer, *Legend of Good Women* (960):

With shippes seven and with no more nauye,
And glad was he to londe for to hye…

It is conceivable that the humorist Gilbert consciously harked back to that.

[2] Here is the second stanza: As office boy I made such a mark
That they gave me the post of a junior clerk.
I served the writs with a smile so bland,
And I copied all the letters in a big round hand.
(CHORUS: He copied all the letters in a big round hand.)
I copied all the letters in a hand so free,
That now I am the ruler of the Queen's Na-vee.
(CHORUS: He copied all the letters in a hand so free,
That now he is the ruler of the Queen's Navee.)

[3] E.g. the Lord Chancellor's song in *lolanthe*: The Law is the true embodiment
Of everything that's excellent.
It has no kind of fault or flaw,
And I, my lords, embody the Law.

The irony is all the more delicious if we happen to know that this librettist had been a barrister. A later stanza, about the Lord Chancellor functioning as guardian of wards in Chancery: And everyone who'd marry a ward
Must come to me for my accord,
And in my court I sit all day
Giving agreeable girls away,
With one for him—and one for he—
And one for you—and one for ye—
And one for thou—and one for thee—
But never, oh never a one for me!

exploits, as in *Pinafore*, the rhyming of pronouns contrary to all normal syntax.

[4] Auber's operatic setting (1856) of the novel by Prévost was destined to lose out eventually to Massenet's (1884)—besides the one in Italian by Puccini (1893).

[5] Gilbert's title came from the quack doctor Dulcamara (= 'Bittersweet', an herb), who sold all kinds of elixirs and other nostrums. Gilbert gave most of the characters the same names as in the opera itself, but he took over just one of Donizetti's airs; the rest he got from various other sources.

[6] Just in recent years, when the Chinese opened their minds wide to the attractions of the Occident, the audiences have reported that *La Traviata* touches their emotions more deeply than any musical drama from their own tradition.

[7] Dr. Dulcamara's spiel, however, is full of *voci sdrucciole* such as these:

Per questo mio specifico,	('Through this my specific,
simpatico, prolifico,	sympathic, prolific,
un uom settuagenario	a man, septuagenarian
e valetudinario	and valetudinarian,
nonno di dieci bamboli ancora diventò.	became again a grandfather of ten babies.')

The music also is somewhat reminiscent of the style of English 'patter' songs: every syllable from *Per* to the end of *valetudinario* is on the pitch of A, with calculated monotony.

[8] From a jolly English operetta it may be a disconcerting shift to *ll Trovatore*, where even the music of love is so somber. But the plot of *Pinafore* was actually derived by Gilbert from *ll Trovatore* with grotesque precision: a hero deprived in infancy of his rightful place among the nobility, through the gypsy Azucena; her comic counterpart, however preposterous, is Little Buttercup.

[9] For the English glosses I resort to the archaic second person singular, because Salvatore Cammarano, Verdi's librettist, makes all the characters consistently *darsi del 'tu'* (as the Italians put it), regardless of any difference in rank or emotional state. He chose not to adhere to the usage of his source, the Spanish play *El Trovador* by Antonio García Gutiérrez, where Leonor at the beginning of the scene addresses De Luna in the second person plural, '¿Me conoceis, Conde?' ('You know me, Count?') but afterwards shifts to the singular: 'Y tu en premio a mi ternura…' ('And thou as a reward for my tenderness…')

[10] Here I show *lágrimé* with two accents. Italian orthography, however, forbids the writing of any accent on *lagrime* or the singular *lagrima*, or any other *voce sdrucciola*. Spanish orthography, on the contrary, requires an accent on lágrima(s). The term *voces esdrújulas* was taken from Italian but modified in accord with Spanish phonology.

[11] Di Luna's part of the duet matches Leonora's metrically, as usual in opera; but since his words—and feelings—clash utterly with hers,

Ah! dell' indegno rendere	('Ah! to render the worthless one's
vorrei peggior la sorte	fate worse I would like
fra mille atroci spasimi	amid a thousand dismal wrenchings
centuplicar sua morte.	to make his death a hundred fold.
Più l'ami e più terribile	The more thou lov'st him, the more terrible
divampa il mio furor!	blazes my wrath!'),

Verdi sets them to a quite different melody. In particular, all three syllables of *rendere*, *spasimi*, and *-ribile* come out on one pitch, middle C. At the close of the duet, when Leonora and Di Luna together sing their last two verses, she repeats the same tune as before; and he sings 'Piu l'ami e piu terribile' exactly an octave lower than her 'Calpesta il mio cadavere', except that the final syllable of his *terribile* drops to D flat while that of *cadavere* rises to F.

Although Verdi's melodramatic effects surpassed than anything earlier, Donizetti had already demonstrated the value of raising the pitch at the end of a *verso sdrucciolo*, as in

Una furtiva lagrima	('A stealthy tear
negli occhi suoi spuntò;	trickled in her eyes;
quelle festose giovani	those cheerful girls
invidiar sembrò:…	she seemed to envy…' *L'Elisir d 'amore*, Act II, scene 8).

[12] That includes two instances of *Mi svena*, the syntactical equivalent of *Svenami*.

[13] For example, when the brief Psalm 117 (116) was translated into Greek, the opening words הללו את יהוה were routinely translated ἀινεῖτε τὸν κύριον 'Praise ye the Lord', with the common noun subsituting for the name itself, just as a Hebrew reader in that era would pronounce the noun אדני [ˀăđonɔ́y] 'my Lord' instead of [y-h(-)w-^{h}] (vowels uncertain). But the closing words of the Psalm, הללו יה, were just translated ἀλληλούϊα. The rule making it a sacrilege to pronounce יהוה did not in principle apply to יה; and the most authoritative pronunciation, [yɔ́h] survived into the medieval (and modern) tradition of

Scripture reading. Neverthelss, in Greek the substitute κύριον (nominative κύριος) was employed to translate יה (Ps. 94 [93]:7, 118 [117]:8, etc.), even when some other form of the verb הלל 'praise' accompanies it: Thus וְעַם נִבְרָא יְהַלֶּל־יָהּ׃ but καὶ λαὸς ὁ κτιζόμενος ἀινέσει τὸν κύριον 'and a people being [now] created will praise the Lord' (Ps. 102 [101]:19). Only in the formula הללו יה was the two-letter divine name (with the Greek pronunciation [ia] or [ya] exempt from being replaced by κυρι-.

[14] It might have encouraged him all the more, if he happened to learn from some Jewish informant that the authentic Hebrew accent belongs at the end, where it gives due prominence to the object of the verb 'praise'.

After my oral presentation in Houston, my friend Barbara Bacz gently reminded me that Handel too (earlier in the 18th century) had dwelt upon the shift of beat in the 'Hallelujah' chorus of *The Messiah*. Now checking the score, I perceive that at different points he gave the beat to each of the four syllables, but much more often it comes on *-lu-* than on any other place. Mozart himself played at least some of the music from this great oratorio and introduced here and there his own contrapuntal additions into the accompaniment. I must leave it to the musicologists to prove whether he was already acquainted with *The Messiah* when he composed *Exsultate, jubilate.* (Handel's restoration of the initial consonant *H-* and the final *-h* was in compliance with the practice of English Protestants since the 16th century, when a number of scholars from England studied Hebrew in Germany. The spelling rules of English suggested to them that many Hebrew words in the English Bible, especially names, could be shown more accurately than had been done in the Latin Vulgate.)

[15] In the repetition it was more graceful to replace the insistent particle *Deh* with the relatively neutral or pathetic *ah*.

[16] The latter option indeed contravenes the phonology of Tuscan otherwise; for the few short words such as *per* and *non* are precursory and never end a sentence or a clause. Nonetheless, since the Iyric poetry in southern Italian dialects, as well as song, was free of such constraints and had been so warmly received throughout Italy, the poets writing in standard Italian arrived at a compromise: in final position too, the *voci piane* can (with some exceptions) drop their final vowel, if preceded by *l*, *m*, *n*, or *r*.

[17] To Bellini alone, Wagner was more favorable than to the other Italians; see Hans-Joachim Bauer, *Richard Wagner Lexikon* (Bergisch Gladbach: Gustav Lübbe, 1988), 68–70. But I have noticed that Bellini too did occasionally sacrifice the proper accentuation, as in Pollione's cavatina in *Norma* (Act I, scene 2):

Udia d'Imene i cantici,	('I heard the chants of Hymen,
vedea fumar gl'incensi,…	I saw the incense smoking,…)

where the normal stress would be on [-d'i-] in *udi(v)a* and on [-d'e-] in *vede(v)a*.

[18] By the ordinary *sandhi* of the Italian language, the noun and adjective *virtù magica* in the first line go so well together as to cancel or reduce the stress that would belong to [-t'u] in any environment other than an adjective such as [m'a-]. No need here to invoke poetic or musical license.

[19] My English gloss keeps the German variation between *Erl-* and *Erlen-*. But for semantic accuracy both *Erlkönig* and *Erlenkönig* should be glossed 'Alderking'; the common noun *Erle* means 'alder'. The credit for this goes to my friend Martin Bidney, professor of English at SUNY-Binghamton.

[20] The weak vowel [-ə] at the end of *ivresse* comes on a high B, startlingly higher than the previous syllables and recalling the mild elevation of [-ə] in *enchanteresse*. The high B

is strengthened by a pause, which marks the sharp contrast between 'folle ivresse' and what follows: 'doux rève'; for the [-ə] of *rêve* is sung on A an octave lower, after the gentle descent of *rê-* from C through B and A down to G, and this A makes a welcome pitch to rest upon.

21 At the end of the second stanza, the score after 'folle ivresse, doux rêve' repeats the previous line 'Charmant souvenir' with the adjective and noun in reverse order, and here too the rhyme with *ravissement* is disregarded.

To me this aria has personal meaning. It was my father's favorite, although he did not understand French. In my early years I heard it often on a Victrola record by Caruso. The old records, along with the phonograph, were abandoned when the family moved; and only once in a while was I reminded of the arias that Caruso had made familiar. But then in November of 1945, when I was sailing home from Naples on a slow troopship, this aria from *Les Pêcheurs de perles* was the only classical piece on the program of recorded music twice a day out on deck. The tenor Richard Crooks sang it, for some reason, in Italian, 'Mi pare udire ancora'; I listened to his rendition dozens of times. Years later I bought a long-playing record of Caruso arias and reacquainted myself with every verse of this one. (From the third verse I have lately taken the password for my e-mail.)

22 I am grateful to my old friend Angela Della Volpe in LACUS and to my dear colleague Peter Cocozzella at SUNY-Binghamton for precious help with some fine points of Italian. It has also been for me a happy occasion to get acquainted with Bruce Borton, professor of music at SUNY-B; he has enabled me—however inexperienced—to record on a cassette tape the excerpts from Gilbert and Sullivan, Verdi, Mozart, Donizetti, Schubert, and Bizet.

* * *

Portions of the scores of works cited, as well as brief recorded passages corresponding to the cited works are on the accompanying CD-ROM. Additional copies of this material may be requested from:

Saul Levin
Department of Classical and Near Eastern Studies
State Univerisity of New York
Binghamton, New York 13902-6000

e-mail: slevin@binghamton.edu

* * *

EXPLICIT/IMPLICIT MODULARITY IN RIGHT HEMISPHERE PRAGMATICS

Jennifer Mortimer
McGill University

introduction. Explicit Memory and Implicit Memory are dissociable and are subserved by different neurofunctional systems. They play complementary roles in supporting those aspects of language that are subserved by the Left Hemisphere: phonological, morphological and syntactic components of language are part of implicit competence, and non-rule governed lexical/semantic knowledge is largely supported by explicit memory. Modularity of linguistic functions in the LH is a theory grounded in more than a century's worth of aphasic studies. In contrast, little research has been conducted regarding the possible dissociability of Right Hemisphere communicative functions (i.e., pragmatic functions: context-dependent and emotional aspects of language comprehension and production). In this paper I will discuss the hypothesis that implicit and explicit memory systems play dissociable and complementary roles in supporting not only components of left-hemisphere linguistic competence but also components of right-hemisphere pragmatic competence. The rationale for this paper will be based on the following points: first, that there is a double dissociation between explicit/implicit memory functions; second, that the explicit/implicit distinction is supported by clinical evidence; third, that different memory systems support different aspects of LH linguistic competence; fourth, that pragmatic competence is lateralized to the RH; and fifth, that there is clinical and experimental evidence to indicate the possibility of a dissociation between explicit and implicit aspects of pragmatics.

1. explicit/implicit memory systems. Explicit knowledge is information that is consciously acquired and consciously used. It is our encyclopedic knowledge of the world—our memories of facts and events. In contrast, implicit competence refers to information that is acquired incidentally, is stored in such a way that it is opaque to introspection, is task-specific, and is used automatically (Paradis 1994). The abilities of implicit memory include the acquisition of skills and habits, non-associative learning, and simple classical conditioning (Squire & Zola-Morgan 1991). Accordingly, access to explicit knowledge is implicit, since the workings of our conscious minds are not consciously controlled.

The neurofunctional model used in this paper to describe the distinction between explicit and implicit memory is based on aspects of Damasio's theory of feature fragments that are compatible with Hebb's cell assembly model (Damasio 1989). In this model, long-term memories are represented in the neo-cortex in the form of networks of locally connected neurons. Individual neurons within the networks are located in areas of the primary and early association cortices that specialize in storing and processing specific types of information: motor, auditory, visual, olfactory, and haptic. The substrate for the convergence of these cells into networks is maintained by the tertiary neo-cortices, as well as by certain structures of the Medial Temporal Lobes (including the hippocampus, the parahippocampal cortex, the entorhinal and perirhinal cortex), the Frontal Lobes, and the Basal Ganglia and Cerebellum. Sensory cortices are interconnected to the hippocampus, and motor and sensory cortices are interconnected to the Basal Ganglia, Cerebellum, and Frontal lobes (Damasio 1989). Clinical studies show that selective impairment to any of these sites typically results in the following impairments: (i) damage to the neo-cortex will destroy the conceptual features/cells which, bound together, constitute memories, thereby rendering these memories inaccessible either in an explicit or an implicit form (ii) damage to the Medial Temporal Lobes (MTL) will result in anterograde and/or retrograde amnesia—the inability to acquire or recall explicit memories—indicating that the MTL is indispensably involved in the consolidation of long-term memories in the neo-cortex (Zola-Morgan & Squire 1990; Alvarez & Squire 1994; Graham & Hodges 1997) (iii) lesions in the region of the frontal cortex known as Broca's area will impair implicit access to linguistic representations, and lesions to the basal ganglia or to the cerebellum will impair use of implicit motor and timing skills (Willingham et al. 1995; Harrington et al. 1998).

In an unimpaired adult speaker, implicit/explicit modularity of linguistic function is represented in the following manner. Lexical-semantic knowledge (arbitrary meanings of words) is considered explicit since this information is learnt and used consciously (we are overtly aware of the meanings of the words we use). This information is maintained in widely distributed areas of the left temporal and parietal lobes. Damage to neuroanatomical substrates, such as the MTL, that are involved in the acquisition of explicit memories, prevents the learning of new lexical items.

Conversely, phonology, morphology, syntax, and some aspects of semantic competence (such as argument structure) are implicit. Among native speakers of a language the use of rules that organize phonemes and morphemes into grammatical utterances is automatic, and the precise form of these rules is impenetrable to conscious analysis. Implicit linguistic competence is subserved in the LH by the ventro-lateral frontal cortex, the basal ganglia, and the cerebellum. These structures are involved in the temporal, sequential organization of information. Damage to these structures typically induces grammatical deficits such as linguistic aprosody, the loss of function words and of rules governing affixation, and in some cases, the inability to comprehend and/or produce syntactically complex sentences (Grossman et al. 1992; Kolk 1998).

2. CLINICAL EVIDENCE FOR DISSOCIABLE MEMORY SYSTEMS. Much of the evidence for a double dissociation between implicit and explicit memory systems is derived from clinical studies. Patients with Alzheimer's Disease, which involves damage to the temporo-parietal neo-cortex and the MTL, are impaired in the explicit recollection of memories and the acquisition of new knowledge (including lexical items and word meanings). Implicit motor and perceptual skills, and grammatical rules, are comparatively unaffected (Ullman et al. 1996). The famous patient H.M., who underwent

bilateral ablation of the MTL, is unable to learn new facts (such as words, his current age), but can still master new motor skills (Milner et al. 1968). Damage to Wernicke's area in the LH typically produces deficits such as anomia, word-finding difficulties, and semantic paraphasias. Grammatical competence is generally spared. Among those patients who show implicit but not explicit memory damage are patients with Parkinson's Disease (linked to the inadequate production of dopamine in the basal ganglia, producing hypokinesia/suppression of movement) and patients with Huntington's Disease (linked to neuronal loss in the inhibitory pathways of the basal ganglia, leading to hyperkinesia/excitation of motor activity). PD and HD patients retain memories of lexical items and of other facts, but incorrectly apply the grammatical rules which govern the use of affixes and function words (Ullman et al. 1996). Damage to the left inferior frontal lobe can result in Broca's aphasia, which is marked by speech apraxia and agrammatism (omission or substitution of grammatical morphemes in obligatory contexts), and an impairment in the skilled organization of phonemes and morphemes into sentences. However, when lexical items can be accessed, they are accessed correctly.

3. RIGHT HEMISPHERE PRAGMATIC COMPETENCE. Interest in the right hemisphere's contribution to communicative competence has waxed and waned since Broca first proposed LH lateralization of language in normals. Hughlings Jackson noted that some aphasics retain the ability to utter swear words under emotional duress, which led him to suggest a distinction between propositional speech (the responsibility of the left hemisphere) and non-propositional speech (the responsibility of the right) (Jackson 1879). The latter refers to 'automatic' speech, including emotional exclamations. Pierre Marie (1906) remarked directly on the preserved pragmatic competence of left hemisphere damaged patients, saying that the 'voluntary politeness of aphasics' was 'additional proof of the conservation in their case of the affective and moral sphere'. These insights into the communicative capabilities of the right hemisphere were more or less ignored until the second half of the twentieth century.

More recently, there has been research to suggest that the right hemisphere has a complementary, rather than a subservient, role in information processing (Sperry 1974; Milner 1980). While the left hemisphere is said to specialize in sequential, context-independent analysis of information, the right hemisphere is holistic, visuospacial, context-dependent, and skilled at analyzing patterns (without analyzing the details of the patterns).

In regards to communication, hemispheric specialization appears to take the form of LH dominance for the linguistic component of communication, and RH dominance for the pragmatic component. The LH's responsibility is to derive meaning from utterances in accordance with a narrow set of inflexibly applied phonological, morphological, and syntactic rules. It is this part of language which is the focus of Chomskian linguistic philosophy. The methodology that drives Chomskian-style linguistic analysis requires examining sentences in isolation; however, ordinary speech does not consist of a string of literal propositions. Consider the following exchange:

A to B: 'I want you to read this book. I am sure you will enjoy it.'
B to A: 'Yes, it looks fascinating.'

Taken by itself, there is only one possible interpretation of B's statement: a literal interpretation. Nevertheless, if the utterance is placed in different contexts, different possible interpretations for B's comment become available. If person A is handing an exciting mystery novel to B, then a literal interpretation of B's reply is appropriate. If person A is handing B a 600-page text on macroeconomics, then a non-literal translation of B's remark is likely to be correct (i.e., B is speaking ironically). In real-life situations, we not only use literal meanings of words and sentences, we also commonly rely on emotional intonation, facial expression, emotional gestures, jokes, sarcasm, and irony to communicate efficiently with each other. While the Left Hemisphere analyzes the literal meaning of speech, the Right Hemisphere appears to be indispensably involved in processing non-literal aspects of communication (Shields 1991; Lorch et al. 1998).

Rigorous testing of RH patients has highlighted language deficits that are consistent with this theory of lateralization of communicative functions. Studies have shown that difficulties with some aspects of prosody (changes in intonation indicating emotional states) generally co-occur with damage to the Right Hemisphere (Borod et al. 1998; Gandour et al. 1995). RHDs may show deficits in understanding facial emotion (Blonder et al. 1991; Borod et al. 1998), affective gestures (Ross 1981), emotional words and sentences (Borod et al. 1992), sarcasm (Kaplan et al. 1990), figurative vs. literal meanings ('He is turning over a new leaf' as opposed to 'He is losing his bad habits') (Van Lancker & Kempler 1987), idiomatic expressions ('Peter kicked the bucket'), jokes (the 'punch-line' is usually context-inappropriate), puns (which rely on a secondary 'context-inappropriate' meaning), and irony (what is said is the opposite of what is meant—a fact which needs to be inferred from the context in which a sentence is spoken). RH patients also seem to have trouble drawing inferences from texts (understanding the theme of a story). In general, RHDs are deficient in grasping non-literal, implied meanings (the term for this condition is *dyshyponoia* (Paradis 1998)).

There are numerous possible models to explain RH dominance in pragmatic abilities. One is that the RH specializes in the expression and recognition of affect (Ross et al. 1981). This theory would account for the apparent RH dominance for paralanguage (emotional prosody, body language, and facial expressions). Another possibility is that the Right Hemisphere possesses a special ability for binding together a number of contextual cues to form meaning. Such an ability would explain how the RH is capable of using contextual information to comprehend non-literal meanings. Recognition of emotion can also be seen as being dependent on inference, in the sense that the perception of emotion requires the flexible interpretation of environmental cues. To understand RH processing, it may be necessary to take into account the contextual domain (discourse, situational or linguistic) from which inferences are made. Equally, the nature of the subject about which inferences are made (i.e., the interlocutor's state of mind) may be relevant.

Unfortunately, although RH dominance for pragmatics is well-supported by clinical studies conducted over the past 20 years, there is little evidence to support any theory that might explain the manner in which pragmatic competence may be represented in the RH. There are few studies, for example, which attempt to correlate RH deficits with intra-hemispheric lesion sites. In fact, studies rarely mention intra-hemispheric site of RH lesion, even though this information is usually available to the authors.

In the following sections I will discuss pragmatic functions in terms of an explicit/implicit memory systems modularity. I will base this programme on studies which have attempted to correlate lesion site with RH deficit. The criteria for determining which memory system supports a particular aspect of pragmatic competence will be the following: (i) explicit pragmatic knowledge must conform to the definition of explicit knowledge (i.e., it must be information learnt and used consciously), and implicit pragmatic skills must conform to the definition of implicit competence (i.e., skills acquired incidentally and used automatically and unconsciously), (ii) there must be clinical or experimental evidence to support classification of pragmatic functions as implicit/explicit. The evidence should show that RH substrates that anatomically correspond with LH substrates supporting explicit linguistic functions in the left hemisphere (i.e., MTL) should be involved in explicit pragmatic functions in the right. RH substrates which support implicit memory (basal ganglia and cerebellum) and/or implicit linguistic competence (inferior frontal lobes or some other focal cortical site) should subserve implicit RH pragmatic functions.

It must be emphasized that the neuroanatomical substrates mentioned above are believed to be indispensably, rather than exclusively, involved in supporting their respective memory systems. For example, damage to neo-cortical areas not directly involved in either memory system may impair a pragmatic function if this cortical area is responsible for processing the sensory information on which the function is based.

4. PROSODY. Prosody consists of speech rhythm (timing and stress) and melody (pitch, tone and intonation). It has at least two functions: one is to support the phonological, morphological, syntactic and lexical elements of language (linguistic prosody), and the other is to convey affective information (emotional prosody) (Hird & Kirsner 1998). The most frequently examined acoustic correlates of emotional prosody are rhythm (expressed as timing and amplitude) and melody (expressed as variations in fundamental frequency (F0)). Lateralization of prosody is generally said to occur along functional lines, with the left hemisphere dominant for linguistic prosody (Gandour et al. 1992a; Moen & Sundet 1996), and the right hemisphere for emotional prosody (Ross 1981). Although a claim has been made for LH specialization in the temporal aspects of prosody, and RH specialization in pitch and tone (F0), this theory is controversial (Gandour et al. 1992b; Niemi 1998).

RH affective prosody deficits (affective aprosody) are notoriously difficult to test without involving other RH pragmatic functions as base tasks. The tasks most commonly used in determining affective prosody deficits require the subject to verbally describe the affective tone of a given utterance, or else to match the affective tone to a picture of a face representing the emotion expressed by the tone (e.g., happiness, sadness, etc.). Both emotional facial recognition and emotional words are believed to be at least partly mediated by the RH. An affective prosodic deficit therefore cannot necessarily be assumed to exist if the determining task involves an aspect of pragmatic communication that has not been controlled for. An affective prosody lateralization study by Ross et al. (1997) has highlighted the seriousness of this problem: Ross et al. found that reduced reliance on verbal processing to determine affective prosody deficits substantially improves LHD but not RHD performance, which indicates that studies of affective prosody lateralization should require as little verbal processing as possible.

5. IS PROSODY IMPLICIT? The appreciation of prosody appears to correspond to the definition of an implicit function: barring exceptional circumstances prosodic skill is used automatically, unconsciously and involuntarily (Cancelliere & Kertesz 1990; Hird & Kirsner 1998). Substrates supporting implicit skills should therefore be involved in the representation of prosody, and damage to these substrates should produce prosodic deficits. Damage to the basal ganglia in PD, HD and stroke patients has been found to co-occur with deficiencies in emotional and linguistic appreciation of prosody (Blonder et al. 1989; Cancelliere & Kertesz 1990; Speedie et al. 1990; Le Dorze et al. 1998).

One would expect that, given that affective prosody is an implicit RH ability, the RH inferior frontal lobes and subcortical basal ganglionic structures in particular should be involved in supporting this function. Cohen et al. (1994) describe the case of a 16-year-old right-handed patient who displayed flat, unanimated intonation after a stroke to the right putamen and the body of the right caudate nucleus. The patient showed 'perfect comprehension of emotional gesturing', but had difficulty in inferring emotion from affect-laden, pre-recorded speech, and was impaired at repeating and expressing affective intonation upon request. Propositional speech was unimpaired. Breitenstein et al. (1998), in a group study of RHD, LHD, PD patients and healthy controls found that patients with anterior RH damage 'showed the strongest impairment in… prosodic emotional recognition' and that PD patients in an advanced stage of the disease showed a similar deficit. The affective prosodic tasks used were two face-matching tasks, a naming task, and a discrimination task. Starkstein et al. (1994) in a group study of RHDs, LHDs, and healthy controls found that the patients with RH lesions involving the basal ganglia and temporal/parietal cortex were the most likely to show aprosody (T/P involvement may reflect the representation of the skill). Patients were asked to identify the emotion in a recorded utterance using a facial emotion matching task. An attempt to control for facial emotion processing was made by giving the subjects a facial emotion recognition task (most, but not all, of the aprosodic subjects performed poorly on this task as well). Infarction of the right thalamus and posterior limb of the internal capsule resulted in aprosody in the case of a 49-year-old woman, as reported by Wolfe & Ross (1987). Spontaneous affective prosody was spared, but affective-prosodic repetition was impaired, as was affective prosody identification in a verbal task. (Although left-handed, the patient was assumed to have a normal lateralization pattern, since comprehension and production of propositional language were intact.) Ross, Harney, deLacoste-Utamsing and Purdy (1981) describe a right-handed patient who possessed normal prosodic comprehension but impaired prosodic affective repetition and extremely poor spontaneous affective prosody after stroke. The site of lesion involved the posterior two thirds of the anterior limb, the entire genu, and the anterior third of the posterior limb of the rostral internal capsule (RH). Again, propositional speech was essentially undisturbed.

6. SARCASM. The interpretation of sarcasm involves knowingly rejecting the literal meaning of an utterance in favour of some non-literal, usually contradictory, meaning. Understanding sarcasm requires the ability to make contextual inferences. For example, of the following statements:

> 'That's a beautiful dress you are wearing' (to someone wearing a faded, ragged dress)
> 'That's a beautiful dress you are wearing' (to someone in expensive, new clothes)

the first has a possible sarcastic interpretation while the second does not. A sarcastic, non-literal interpretation of the first remark is cued by the discrepancy between the literal meaning of the remark and the meaning that would seem appropriate given the context.

The appreciation of sarcasm is at least partly an explicit function. Its explicitness is evident from the fact that we can consciously identify the element in the situational context that makes the literal interpretation impossible (in the first example, the fact that the dress was faded and ragged). Moreover, in everyday conversation the comprehension of sarcastic statements is not necessarily automatic: occasionally, the interpretation of a sarcastic comment requires an explicit explanation from the speaker.

Deficits in processing sarcasm have been linked to right hemisphere damage (Brownell et al. 1986) and frontal lobe damage (Alexander et al. 1989). McDonald and Pearce (1996) have found loss of ability to interpret sarcastic remarks in patients with frontal lobe and subcortical damage (8 of 10 subjects with bi-lateral damage). Their test required the subjects to read a short dialogue and then answer questions about the exchange. For example:

Mark:	'What a great football game'.
Wayne:	'Sorry I made you come'.
Question:	Did Mark think the game was good/bad?

Seven of the ten subjects were also given an affective prosody recognition task. Although some of their subjects were impaired on this task, there was no correlation between their ability to detect emotional tone and their ability to recognize sarcasm.

These results indicate that the interpretation of sarcasm is not dependent on correctly processing the prosodic tone in which the remark is uttered. McDonald and Pearce also evaluated their subjects' performance on the Wisconsin Card Sorting Task. Here there was a direct correlation between the results on the WCST and results on the sarcasm interpretation task. Performance on the WCST is usually disrupted after lesion to the dorsolateral frontal cortex, which supports the ability to consciously change one's reactions in the face of changing contextual cues (Petrides 1989). Damage to the dorsolateral frontal cortex may cause difficulty in switching from a literal to a non-literal interpretation of an utterance. McDonald and Pearce's (1996) results are consistent with a proposal by Alexander et al. (1989) that lesions to the RH dorsolateral frontal cortex may cause disturbances in the appreciation of sarcasm, irony, and inference.

7. METAPHORS. A metaphor is a descriptive term given to an object or action to which it is imaginatively, but not literally, related. As in the case of the appreciation of sarcasm, the comprehension of metaphors necessitates rejecting a literal interpretation in favour of a figurative one. The correct interpretation of a metaphor is derivable from the context (intra- or extra-sentential) in which it is produced. For example, in the sentence: 'The lady's jewels were bursting stars' the non-literal meaning of the metaphor is understood through comparing the properties common to the term *bursting stars* and to the other noun in the sentence, *jewels*. Of the network of conceptual features mapped to the lexical items *bursting stars* (i.e., 'bright', 'sparkling', 'gaseous', 'enormous') only those features that are shared with the lexical item *jewel* are understood to be relevant in this context (i.e., 'bright' and 'sparkling'). Explicit memory is involved in supporting the appreciation of metaphors, since the ability improves with instruction (Paradis class notes 1999). The inferencing process used to understand metaphors can therefore be explicitly taught.

Metaphor interpretation was investigated by Bottini et al. (1995) in an experimental PET study that involved scanning both hemispheres of students participating in a metaphor-judgement task. The students were presented with sentences containing metaphors, and were asked to indicate whether they found the metaphor plausible or implausible. For example:

'The investors were squirrels collecting nuts.'	(Plausible)
'The investors were trams.'	(Implausible)

This task condition was compared to a literal sentence processing condition in which students judged the sentence's plausibility at the literal level (i.e., 'The boy used stones as paperweights'—plausible). The results of the comparison showed activation in the RH middle temporal gyrus, the RH precuneus (which is reportedly involved in long term memory functions such as the retrieval of words from memory (Bottini et al. 1995)), the RH anterior cingulate (which appears to play a role in conscious, controlled processing (Dehaene et al. 1997)), and the RH dorsolateral frontal cortex, which, as mentioned in relation to sarcasm, is involved in processing conditional cues.

8. PREDICTIVE MODELS OF RH FUNCTIONS. This paper has dealt exclusively with prosody, sarcasm, and metaphor. Other aspects of pragmatics were not discussed because of a lack of clinical and experimental evidence regarding their relations to explicit/implicit memory. It is nonetheless possible, using the definitions of explicit and implicit memory, to predict the roles of these memory systems in sustaining pragmatic functions not discussed in this paper. For example, the appreciation and expression of affective body language (i.e., gaze and gestures) and of facial expression is largely automatic/involuntary, which suggests implicit memory involvement in their representations. Some of the aprosodics in the case studies reviewed were impaired in the interpretation of facial expression (Starkstein et al. 1994), however, some were not, indicating that although the interpretation of facial expression and the recognition of affective prosody may both be implicit pragmatic functions, they are also neurofunctionally dissociable. Lexical-semantic information is mediated to a certain extent by explicit memory in the LH, which suggests that RH explicit memory substrates (i.e., MTL/TP cortex) are involved in lexical pragmatic functions (jokes, irony, figurative meanings). The deliberate inferencing of non-literal meanings from context may be partially explicit, as McDonald and Pearce's (1996) study indicates. The RH dorsolateral frontal cortex may also be involved in pragmatic functions that rely on conscious lexical inferrencing processes (as seen in Bottini et al.'s 1995 study of metaphor comprehension).

9. CONCLUSION. The intent of this paper was to a) demonstrate a lack of available information regarding the manner in which RH pragmatic function is represented, and b) provide arguments in terms of an application of implicit/explicit memory theory as to why this dearth should be addressed. In this paper I have argued in favour of an RH lateralized dissociation between implicit (e.g., affective prosody) and explicit (e.g., metaphor processing) pragmatic function. To the extent that appropriate studies were available for review, the data do support an apparent implicit/explicit memory dissociation of RH pragmatics. However, this hypothesis needs to be supplemented by further studies.

REFERENCES

ALEXANDER, MICHAEL D., FRANK BENSON & DONALD T. STUSS. 1989. Frontal language. *Brain and language* 37:656–91.

ALVAREZ, PABLO & LARRY R. SQUIRE. 1994. Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Science* 91:7041–45.

BLONDER, LEE XENAKIS, RAQUEL E. GUR & RUBEN C. GUR. 1989. The effects of right and left Hemiparkinsonism on prosody. *Brain and language* 36:193–207.

BLONDER, LEE XENAKIS, DAWN BOWERS & KENNETH M. HEILMAN. 1991. The role of the right hemisphere in emotional communication. *Brain* 114:1115–27.

Borod, Joan, Fani Andelman, Loraine K. Obler, James R. Tweedy & Joan Welkowitz. 1992. Right hemisphere specialization for the identification of emotional words and sentences: evidence from stroke patients. *Neuropsychologia* 30(9):827-44.

Borod, Joan, Barbara A. Cicero, Loraine K. Obler, Joan Welkowitz, Hulya M. Erhan, Cornelia Santschi, Ilana S. Grunwald, Reto Agosti & John R. Walen. 1998. Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology* 12(3):446–58.

Bottini, G., R. Corcoran, R. Sterzi, E. Paulesu, P. Schenone, P. Scarpa, R. Frackowiak & C. Frith. 1995. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain* 117:1241–53.

Breitenstein, Caterina, Irene Daum & Hermann Ackermann. 1998. Emotional processing following cortical and subcortical brain damage: Contribution of the fronto-striatal circuitry. *Behavioural neurology* 11:29–42.

Brownell, Hiram, Heather H. Potter, Amy M. Bihrle & Howard Gardner. 1986. Inference deficits in right brain-damaged patients. *Brain and language* 27:310–21.

Cancelliere, Andy E. B. & Andrew Kertesz. 1990. Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and cognition* 13:133–47.

Cohen, Morris, Cynthia A. Riccio & Ann Marie Flannery. 1994. Expressive aprosodia following stroke to the right basal ganglia: A case report. *Neuropsychology* 8:242–45.

Damasio, Antonio R. 1989. Time-locked multi-regional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition* 33:25–62.

Dehaene, Stanislaus, Emmanuel Dupoux, Jacques Mehler, Laurent Cohen, Eraldo Paulesu, Daniela Perani, Pierre-Francois van de Moortele, Stephane Lérircy & Denis Le Bihan. 1997. Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8:3809–15.

Gandour, Jack, Suvit Ponglorpisit, Fuangfa Khunadorn, Sumalee Dechongkit, Prasert Boongird & Rachanee Boonklam. 1992a. Timing characteristics of speech after brain damage: vowel length in Thai. *Brain and language* 43:337–45.

——— et al. 1992b. Lexical tones in Thai after unilateral brain damage. *Brain and language* 43:275-307.

Gandour, Jack, Jennifer Larsen, Sumalee Dechongkit, Suvit Ponglorpisit & Fuangfa Khunadorn. 1995. Speech prosody in affective contexts in Thai patients with right hemisphere lesions. *Brain and language* 51:422–43.

Graham, Kim S. & John R. Hodges. 1997. Differentiating the roles of the hippocampal complex and the neocortex in long term memory storage: Evidence from the study of Semantic Dementia and Alzheimer's Disease. *Neuropsychology* 11:77–89.

GROSSMAN, MURRAY, SUSAN CARVELL, MATTHEW B. STERN, STEPHEN GOLLOMP & HOWARD I. HURTIG. 1992. Sentence comprehension in Parkinson's Disease: The role of attention and memory. *Brain and language* 42:347–84.

HARRINGTON, DEBORAH L., KATHLEEN L. HAALAND & N. HERMANOWICZ. 1998. Temporal processing in the basal ganglia. *Neuropsychology* 12:3–12.

HIRD, K. & KIM KIRSNER. 1998. Control processes in prosody. In *Implicit and explicit mental processes*, ed. by Kim Kirsner, 201–18. Mahwah NJ: Lawrence Erlbaum Associates.

JACKSON, JOHN HUGHLINGS. 1879. On affections of speech from disease of the brain. Reprinted in *Reader in the history of aphasia*, ed. by Paul Eling, 1994. Amsterdam: John Benjamins.

KAPLAN, JOAN A., HIRAM H. BROWNELL, JANET R. JACOBS & HOWARD GARDNER. 1990. The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain and language* 38:315–33.

KOLK, HERMAN. 1998. Disorders of syntax in aphasia: Linguistic-descriptive and processing approaches. In *Handbook of neurolinguistics*, ed. by Brigitte Stemmer & Harry Whitaker, 249–60. London: Academic Press.

LE DORZE, GUYLAINE, JOHN RYALLS, CHRISTINE BRASSARD, NANCY BOULANGER & DANIELLE RATTÉ. 1998. A comparison of the prosodic characteristics of the speech of people with Parkinson's Disease and Friedreich's Ataxia with neurologically normal speakers. *Folia phoniatrica et logopædica* 50:1–9.

LORCH, MARJORIE P., JOAN C. BOROD & ELISSA KOFF. 1998. The role of emotion in the linguistic and pragmatic aspects of aphasic performance. In *Pragmatics in neurogenic communication disorders*, ed. by Michel Paradis, 103–118. Oxford: Elsevier.

MARIE, PIERRE. 1906. Translated from 'La troisième circonvulation frontale gauche ne joue aucun role spécial dans la fonction de langage'. Reprinted in *Reader in the history of aphasia*, ed. by Paul Eling, 1994. Amsterdam: John Benjamins.

MCDONALD, SKYE & SAMANTHA PEARCE. 1996. Clinical insights into pragmatic theory: frontal lobe deficits and sarcasm. *Brain and language* 53:81–104.

MILNER, BRENDA. 1980. Complementary functional specialization of the human cerebral hemispheres. In *Nerve cells, transmitters, and behaviour*, ed. by R. Levi-Montalcini. Rome: Pontifica Academia Scientarium.

———, SUZANNE CORKIN & H.-L. TEUBER. 1968. Further analysis of hippocampal amnesic syndrome: 14-year follow-up study of H.M. *Neuropsychologia* 6:215–34.

MOEN, INGER & KJETIL SUNDET. 1996. Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and language* 53:267–81.

NIEMI, JUSSI. 1998. Modularity of prosody: Autonomy of phonological quantity and intonation in aphasia. *Brain and language* 61:45–53.

PARADIS, MICHEL. 1994. Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In *Implicit and explicit learning of languages*, ed. by Nick C. Ellis, 393–419. London: Academic Press.

———. 1998. The other side of language: Pragmatic competence. In *Pragmatics in neurogenic communication disorders*, ed. by Michel Paradis, 1–10. Oxford: Elsevier.

Petrides, Michael. 1989. Frontal lobes and memory. In *Handbook of neuropsychology*, ed. by F. Boller & J. Grafman, 75-90. Oxford: Elsevier.

Ross, Elliott D. 1981. The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of neurology* 38:561–69.

———, John H. Harney, Christine deLacoste-Utamsing & Phillip D. Purdy. 1981. How the brain integrates affective and propositional language into a unified behavioral function. *Archives of neurology* 38:745–48.

———, Robin D. Thompson & Joseph Yenkosky. 1997. Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and language* 56:27–54.

Shields, Jane. 1991. "Semantic-pragmatic disorder: a right hemisphere syndrome?" British journal of disorders of communication, 26:383-92.

Speedie, Lynn. J., Nancy Brake, Susan Folstein, Dawn Bowers & Kenneth M. Heilman. 1990. Comprehension of prosody in Huntington's disease. *Journal of neurology, neurosurgery and psychiatry* 53:607–10.

Sperry, R. 1974. Lateral specialization in the surgically separated hemispheres. *In Neurosciences: Third study program*, ed. by F. O. Schmidt & F. G. Worden. Cambridge MA: MIT Press.

Squire, Larry R. & Stuart Zola-Morgan. 1991. The medial temporal lobe memory system. *Science* 253:1380–86.

Starkstein, S.E., J. P. Federoff, T. R. Price, R. Leiguarda & R. Robinson. 1994. Neuropsychological and neuroradiologic correlates of emotional prosody comprehension. *Neurology* 44:515–22.

Ullman, Michael T., Suzanne Corkin, Marie Coppola, Gregory Hickok, John H. Growdon, Walter Koroshetz & Steven Pinker. 1996. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of cognitive neuroscience* 9(2):266–76.

Van Lancker, Diana R. & Daniel Kempler. 1987. Comprehension of familiar phrases by left- but not by right-hemisphere-damaged patients. *Brain and language* 32:265–77.

Willingham, Daniel B., Walter J. Koroshetz & Evelyn W. Peterson. 1996. Motor skills have diverse neural bases: spared and impaired skill acquisition in Huntington's Disease. *Neuropsychology* 10:315–21.

Wolfe, Gil I. & Elliott D. Ross. 1987. Sensory aprosodia with left hemiparesis from subcortical infarction. *Archives of neurology* 44:668–71.

Zola-Morgan, Stuart & Larry R. Squire. 1990. The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science* 250:288–90.

* * *

THE AMPLIFICATION CLAUSE ROOT IN SAN MIGUEL EL GRANDE MIXTEC

Evelyn G. Pike
SIL International

San Miguel el Grande Mixtec (SMM) is spoken in the western highlands of the State of Oaxaca, Mexico. The data are taken from more than fifty-five discourses including narrative, procedural, behavioral, and expository types having to do with the beginnings of various phenomena, the spirit world, causes and cures of various illnesses, causes of various anomalies of nature, customs, and animal and people stories. These were collected some 55 to 65 years ago. Many of these were told and then written by one of the leading storytellers of the village, Narciso Merecías. The goal of this paper is to identify and uniquely characterize the form and function of an interesting grammatical structure in SMM.

1. overview. Very early in my time with the people of San Miguel el Grande, I heard and learned a statement about a donkey hurrying down a trail. It was:

^{3}*Ndaba*	*-tə̄*	*cuahān*	*-tə̄.*
jumpPOT	-3AML	goPOT	-3AML[1]

‘It goes bouncing along.’

This seemed to me to describe most accurately what the donkey was doing. Then I began to notice more of such constructions that primarily, but not exclusively, use verbs of directional motion. Also this P (simple predicate) is followed by a S (subject) which is coreferential with the S following the second P. The two Ps are never the same, but each amplifies the meaning of the other, and I never heard a pause between the two PSs. Because of these factors, I have judged the combination to be a special sequence, which occurs in the same place in the nucleus of a clause that a simple PS group does. So I have labeled it Amplification Clause Root (AmpliClRt). This structure may have up to three PS groups.

A second factor supporting the unity of the AmpliClRt has come to attention in the last several years. Some of the longer texts that were first written by Narciso Merecías were recently edited by the very knowledgeable Angel Merecías Sánchez by adding the very necessary symbols for tone, glottal stop, length of vowel, and nasalization. The editing also adds commas indicating phonological breaks, but *no* periods, which would indicate sentence breaks. Importantly for my analysis, there are no commas between the PSs of what I’ve identified as the AmpliClRt. These data give evidence that mother-tongue speakers recognize the phonological unity of that structure.

Section 2 presents data of clauses (ClRt +/- ClMargins of time, location, manner, etc.) in which are the more frequently occurring clause roots (P + S + any obligatory or optional argument). Only with this information is there a basis to judge that the amplification structure does occur where other clause roots occur.

Section 3 presents three levels of sentence-independency in relation to Angel's commas. These sentence features need to be known in order to determine whether the amplification structure, without commas, best fits in the nucleus of a sentence (normally filled by a clause) rather than in the nucleus of a clause. The data support the latter.

Section 4 presents examples of the content and distribution of the 76 AmpliCl-Roots found in the corpus of more than fifty-five discourses. It shows that, indeed, the AmpliClRt occurs in the nucleus of the clause where other clause roots occur.

Section 5 presents examples that closely resemble the AmpliClRt, but are shown to be contrastive with it.

Section 6 presents the distribution of aspect in each VP of the AmpliClRt. However, such features have yet to be related to aspect and time relations of the larger context.

Section 7 summarizes some of the interesting features of the AmpliClRt.

2. MORE FREQUENT CLAUSE STRUCTURE. In most of the nuclear clause structures the order of constituents is rigidly VSO. In this paper the V will be identified as P (simple predicate). P, as a slot, is filled by the class Verb Phrase (VP) the minimum of which is a single verb. The verb consists of a single verb root or up to a complex of verb roots, augmentable by several prefixes. The VP optionally includes some limiting particles and adverbial phrases following the verb, and is optionally preceded by one of two aspect proclitics, which in turn may be preceded by various verb modifiers. In parentheses, following the English translation of the examples, is the number of the discourse and place in the discourse from which the example is taken. For a statement of grammatical levels of structure used here, see Pike and Pike (1982 and 1983).

Single VRt (bold face) is a minimal VP followed by S (with a hyphen in the Mixtec indicating a border between morphemes within a word, and a dot between English words indicating the meaning of a single Mixtec morpheme):

(1)	…	*te*	***quee***	*ndúú*	*-ni*	*máá**	*-yó*…*
		and	eatPOT	two	-just	just.that.one	-12

'… just the two of us will eat…' (25.5)

Example (2) begins with a frequently occurring sentence connector (SC), and then is followed by an independent clause which begins with a VP (bold face) in which the nucleus is a close-knit verb stem 'be.happy' followed by S, and then a clause margin with the role of reason. The part which follows *jā** is a transitive clause root with its VP in bold face :

(2) ^{1}*Te yúan te* ***nī*** ***cā-cu-səə̄**** ***xáān* iní*** *-de jí* -náhan*
and then and cmp pl-be-happy very stomach -3fm with -group
*-de, jā** ***ní*** ***ca -nihīn**** ***xáān**** *-de nujíí, jíín* parque.*
-3fm that cmp pl-obtain very -3mf weapon with ammunition
'Then they all were very happy, that they had obtained many weapons and ammunition.' (54.589–91)

In (3) the VP is followed by S (*-ña** is an enclitic variant of *naha* 'woman' and *-ún* is an enclitic variant of *yúan* 'that.one') and Argument (slot) as Undergoer (role), (Arg-U):

(3) *…te* ***nī*** ***quihin*** *-ña* chúun -ña* -ún…*
…and cmp take -3ff chicken -3ff -that.one
'…she caught that specific chicken of hers…' (25.7)

Example (4) shows a more complex VP:

(4) ^{1}*Yúan-na te* ***nī*** ***cā -na -caca yátá*** *-de jí* -náhan -de.*
then-just and cmp pl -rep -walk backwards -3fm with -group -3fm
'Then all these retreated again.' (54.450)

In the following examples the nucleus of the clause is in bold face. In addition to the Arg-U, Mixtec can also include an Arg-Sc (Scope) which is the direction of the action. Following the clause nucleus (a slot), optionally, are clause margins (slots) with roles such as time, manner, purpose, location, etc. Slot and role are two of the four features of the tagmeme, in our model of Tagmemics. The other two features are the filler class (VRt, VP, NP, LocPh, TClRt, etc.) and cohesion (agreement constraints). Cf. Pike and Pike (1982 and 1983).

Example (5) shows ClRt: P S U + ClMar-Purpose:

(5) *súan* ***x**** ***-ndóo*** ***-de*** ***əən vansada,*** * *cá -i*
thus cs -remaincnt/pot -3fm one outpost cnt pl -?
-ndīto ichi Carranza*
-take.care.of road Carranza
'…in this way they were placing personnel at an outpost in order to throw back the Carrancistas.' (54.455–56)

In example (6) SC is followed by ClRt (bold face): P S U Sc + ClMar-Loc:

(6) ^{1}*Jííān* ***te*** ***nī*** ***nda -nihin -de*** ***batéa*** ***yúan*** ***səquə̄**** ***-de*** *yúan.*
then and CMP raise-obtain -3FM trough there on.top.of -3FM there
'Then, there, he lifted up that trough on top of himself.' (54.696–97)

Examples (1)–(6) present the structure of the most frequently occurring clauses in SMM, showing the sequence of P, S, Arg, and ClMargins. If the PSs of the amplification structure can each have their own arguments and clause margins without a comma before a following P, then, probably considering each a separate clause would be the better analysis.

3. INDEPENDENCY OF SENTENCE. In SMM the S, any Arg, or any ClMar may be fronted, in which case these fronted elements are sentence margins (SentMar). The fronted unit is in special prominence according to discourse features. Other sentence constituents are sentence connectors (SC) which precede the fronted elements. Depending on the occurrence of these sentence constitutents, we propose at least three levels of sentence independency within a discourse. The most independent sentences are those introduced by fronted units or by SCs followed by fronted units. Such sentences are indicated by a superscripted 1 (^{1}). The next level of independency are sentences introduced by SCs, but no fronted units. Such sentences are indicated by a superscripted 2 (^{2}). The least independent sentences are those with no SCs and no fronted unit, hence they begin with the Predicate. They are indicated by a superscripted 3 (^{3}). These latter are, with very rare exceptions, a re-statement of what has already been said.

This information is relevant to determine if the amplification structure not only has the internal structure of a clause, but also the distribution within a sentence that a clause has.

The sentence analysis above has been done with the data of Text 54, which is judged to have more than 500 sentences. This is a narrative which was first written by Narciso Merecías. It tells of the involvement of the men of San Miguel el Grande with the government army of the State of Oaxaca fighting the revolutionaries Venustiano Carranza, Francisco Villa, and Emiliano Zapata, before the actual commencement of the Mexican Revolution in 1910. This is one of the texts which was edited by Angel Merecías Sánchez. In his editing he indicated phonological breaks with commas, but no periods at all to indicate grammatical sentences. That problem, then, gave me the stimulus to examine closely what I consider to be sentence features of SMM. Every period that I've introduced in the SMM text replaces a comma. All other commas are as Angel wrote them, and which I judge not to indicate sentence breaks. The data here show that the amplification sequence fits better as a clause root in the nucleus of the clause rather than as a clause in the nucleus of a sentence.

In (7) the P and S of the clause root are in bold face. Note the SC *te* plus the fronting of S, which is a NP coreferenced in the normal position by the pronoun *-de* (enclitic variant). Two LocPhs ('at the designated place' and 'at the outpost', each preceded by a comma) follow:

(7)	^{1}_Te_	_sava_	_-ga_	_fuerzá_	**_nī_**	**_cā_**	**_-s*_**	**_-quéndōo_**	**_-de,_**	_nū*_	_ní_
	and	some/half	-again	force	CMP	PL	-CS	-remain	-3FM	face	CMP

tetahān	_-de,_	_nū*_	_cóo_	_vansadá._
assign.work	-3FM	face	beCNT/POT	outpost

(7) ‘The rest of the forces he required to stay put at the designated place, at the outpost.’ (54.438–40)

In (8) the SC *núsáá te* and fronted time phrase *vita ñúhni* are followed by a clause root:

(8) ^{1}*Núsáá te vita ñúhni* ***ná* cuáha -rī nombramiento -de.***
then and now immediately INSIS givePOT -1I assignment -3FM
‘Then right now, I’ll give him his commission.’ (54.103)

In (9) note the SC *te* preceding the fronted TmCl and the *te* that introduces the clause nucleus (bold face). Also note in the TmCl that modifies ‘paper’ is an independent clause, which in a different environment would be a type ^{3} sentence, as also is the modification of Mr. Soria translated as ‘who would come with it’:

(9) ^{1}*Te nū* ní jīnu tutū yúan nī sáha -de, te* ***nī***
and when CMP finish paper there CMP do -3FM and CMP
jáha -de nuū* súu chāa Soria vāi -de jíín*.
give -3FM face DEF.ITEM man Soria comePOT -3FM with
‘When he finished that paper which he had done, he gave it to Mr. Soria who would come with it.’ (54.110–11)

In (10) the SC *te* appears with no fronted element:

(10) ^{2}*Te nī jaā -de yúan, nuū* cháa * cúu Gobernador.*
and CMP arrive -3FM there face man CNT be governor
‘He arrived there, in the presence of the man who was the Governor.’ (54.16–17)

In (11) the least independent sentence, beginning with P, summarizes the preceding sentences:

(11) ^{3}*Vāi -de nə́ə́ tácá calle yúan.*
comePOT -3FM the.whole all street there
‘He was coming through all those streets. (54.701)

4. AMPLIFICATION CLAUSE ROOT. The PS groups in the amplification structure of SMM are a feature of the Otomanguean language family. In languages such as Trique, Chatino, Otomi, and other Mixtec languages these structures are described as occurring at a higher level rather than in the nucleus of the clause (Pride 1965 and Alexander 1988). Longacre 1966 describes them as constituents of the sentence.

In the examples presented so far are given the most frequent PS sequences in the nucleus of the clause, along with some other features of the SMM sentence. The data are taken from more than 1,500 sentences in texts. Almost all sentences have in their nucleus an independent clause. There are also many independent clauses which are modifiers, particularly in a NP (cf. (9)) or introduced by a dependent marker (cf. (2)). All these are sources from which at least 76 examples of the amplification structure have been collected. The most frequent verb roots involved are directional motion ones. There are 52 examples (68%) in which all the verbs are of this type. The VRt *quee* 'leave' occurs 23 times (30%), but only once as the second P of the group. As the first P, it occurs 11 times (14%) with *cuahān* 'go' as the second P. This is the most frequently occurring pair of verbs. There are 16 (21%) examples in which there are no directional motion verbs. There are 10 (13%) examples in which one verb is a directional motion verb. The nondirectional motion verbs have the translations of do, say, sleep, fear, think, sing, yell, register (as for a marriage license), carry.by.hand, remain, win, throw.out, be.sweet, be.responsible.for, be.situated.on, be.carrying, be.hidden, be.located. The Ps of the AmpliClRt, in this section, are in bold face. These data continue to support the claim that the amplification sequence is a clause root, especially since no comma nor Arg nor ClMar occurs between the PSs.

In (12) three PSs include *quee* and *cuahān*:

(12) ^{1}*Yúan -na te* ***nī*** ***quee*** *máá** *cháa-ún* ***cuahān**-de*
then -just and CMP leave that.one man-that goPOT -3FM
cuā -nohon *-de vehe -de.*
goPOT -go.home -3FM house -3FM
'Then that man went away going home to his house.' (21.9)

In (13) three PSs appear with only two directional motion verbs:

(13) ^{2}*Te* ***nī*** ***quihin**-ní* * ***cándáhá*** *-ní* ***vāi*** *-ní jíín*.*
and CMP take -2F CNT carry.by.hand -2F comePOT -2F with
'You took it and came carrying it by hand.' (54.242)

In (14) two PSs appear with the most frequently occurring verb pairs *quee* followed by *cuahān* varying to the allomorph *jahān*. Arg-Sc follows:

(14) ^{1}*Chāa yúan* ***nī*** ***quee** -de* ***nī*** ***jahān**-de, ñuū Nunduá…*
man there CMP leave -3FM CMP go -3FM town Oaxaca.City…
'That man, he left, going to Oaxaca City …' (54.5)

In (15) the two Ps are not verbs of directional motion. Note the order of the Ps: drink is first, then eat. Then note the order of the Us (not according to the order of their verbs): food, coffee, bread. Only coffee, which is in the middle, is what is drunk; the other two are what are eaten. If the PSs were not a unit, but rather a sequence of two clauses then coffee would have occurred following the verb drink. Also the feminine form is used as subject even though men are also eating and drinking. This is because women are responsible for providing the food:

(15) ^{2}*Te* ***nī*** ***cā-jihi*** *ndénúní* -ñá* ***nī*** ***cā-yee**** *-ñá ndeyu jíín**
and CMP PL-drink the.three -3FF CMP PL-eat -3FF food with
cájuéē jíín státilá.*
coffee with bread
'And the three couples ate the food, coffee, and bread.' (21.55)

In (16) two other directional motion verbs are followed by a time-location phrase. Cf. (11):

(16) ^{3}***Nī ndāvə*** *-de* ***vāi*** *-de, ondē nū* ní ndu -ñaa.*
CMP to.enter.again -3FM comePOT -3FM until face CMP become -darkness
'He re-entered (the streets) coming to where it was dark.' (54.704–05a)

Example (17) shows different set of verbs with Sc:

(17) ^{2}*Te* ***nī*** ***quee*** *-de* ***nī*** ***ja*** ***-cachī*** *-de nuū* lamú.*
and CMP leave-3MF CMP arrive-say -3MF face owner
'He went and told the owner.' (16.25)

For (18) note that the two Ps (verbs of directional motion) are *quee* 'leave' and *nohon* 'go.home'. The first S is expressed by a long NP including both a woman and two males, and the second S, coreferential with the first S, uses the feminine pronoun even though two men are involved. Note the plural prefix in the second verb. Once again *-ña** occurs because it's the woman's responsibility for getting the family home. Arg-Sc follows the PSs:

(18) ^{1}*Yúan -na te* ***nī*** ***quee*** *ñahan yúan jíín* cháa yúan jíí* sé*
then -just and CMP leave woman that with man that with child
*-yíí- ña** ***cá-nohon*** *-ña* véhe -ña**
-male -3FF PL -go.home -3FF house -3FF
'Then that woman, that man, and their son went back home to their house.' (21.63)

5. COUNTEREXAMPLES. There are complex VPs which might at first glance seem to be AmpliClRts, but are separated by a comma, or have two Ps but just one S, or have the same P.

In (19) a comma separates the two PSs.

(19) ^{1}*Yúan te nī jácunchaā -de, nī cavā -de ñayəvə*.*
then and CMP begin -3FM CMP twist -3FM person
'Then he began to gather the people together.' (54.405–06)

A comma separates two PSs in (20): I analyze it as two sentences because the second sentence reiterates what has already been said. I have replaced the comma with a period:

(20) [2]*Te nī cā-cu-tútú* -i jí* -náhan-i.* [3]*Nī ca-hīin -i.*
and CMP PL-be-assemble-3I with-group-3I CMP PL-be.located-3I
'They all came together. They all were there.' (54.407–08)

Example (21) is a complex VP followed by one S:

(21) [2]*Te nī ndəhə nī cā-sáha tūha -de máá* -de jí**
and CMP to.end CMP PL-do prepare-3FM just.that.one-3FM with
náhan-de.
group-3FM
'They had finished their precautionary preparations.' (54.613)

Example (22) is another complex VP in which the two verb roots are the same, each with a different modifier. These units are followed by *iní* which is obligatory to both verb roots to communicate 'to.decide' rather than 'to.remove', which it would mean without *iní*. There is only one S:

(22) *…ní tava* ndáā nī tava* cuə́tə* iní -ná…*
…CMP decide true CMP decide absolute stomach-1F
'…I have decided truly and absolutely …' (21.28)

Example (23) is more ambiguous as to whether what I've considered to be in the nucleus of the clause is rather in the nucleus of the sentence, as some have claimed for other Otomangean languages. In that case yúan would be a clause margin of location and the following PS would be the second PS of an Amplification Clause . However, I judge that the second PS is modifying yúan to form a LocP and not a second PS of an amplification structure. This results in having just one PS in the clause nucleus:

(23) [3]* *cá-i-ndātu-de yúan* * *cá-hīin-de.*
CNT PL-?-wait -3FM there CNT PL-be? -3FM
'They were all waiting there where they were.' (54.419)

6. ASPECT SEQUENCES. The combination of aspects in the AmpliClRt is complex. Here will be presented some of the data regarding the aspect combinations in the two SP sequences.

Aspect is indicated in the VP. Time within a discourse is indicated by clause margins of time or by time units at higher grammatical levels. There are three aspects in SMM: 1) potential (POT), 2) completive (CMP), and 3) continuative (CNT). POT is the basic form of the verb root. Each aspect may occur with any time unit. POT occurs most frequently with a future time. CMP is indicated by the morpheme *nī* (varying to *ní*) immediately preceding the verb in the VP. It occurs most frequently with a past time unit. CNT is indicated, in the same position as ní, by the continuative aspect morpheme which, in SMM, has no consonants nor vowels (but does in some other Mixtec languages). It has only the tone-changing features of an arbitrary set of morphemes (B class) that changes the tone sequence of the following morpheme, if that sequence can be changed. That continuative morpheme is indicated by an asterisk on the line (*). Other morphemes in that arbitrary B class are marked by a raised asterisk following the morpheme (morpheme*). The aspect of a VP is ambiguous (CNT/POT) when the verb has a changed tone sequence (but with no variant having segmental phoneme changes) and is preceded by a B class morpheme (with consonants and/or vowels). CNT occurs most frequently with a current time indicator. Cf. Pike (1948:77–94) for details of the morphosyntactic tone rules of SMM.

The occurrence of aspect and time units in a text are very important in understanding the structure and the meaning of that text. We begin here by examining the distribution of aspect in the 72 examples of the AmpliClRt taken from the entire corpus. This is seen in the first column of the Table 1 (overleaf). In the second column are the distribution features of aspect in the 36 examples of the AmpliClRt taken from Text 54. The reason for this contrast is that Text 54 is a narrative and its basic aspect is CMP. I haven't studied in any detail the genres and basic aspect of most of the other discourses; it is a heterogeneous group. However, Text 54 is a narrative, hence more homogeneous overall, although embedded in it are sections of other types.

The highest frequency of the nine possible combinations of aspect are CMP + POT: 21% for the total corpus and 31% for Text 54. This higher percentage for Text 54 might possibly have some relation to its basic CMP aspect. The percentage for CMP in one or both PSs is 58% for the total corpus, and 64% for Text 54. Such a difference isn't as great as I would have expected.

In Text 54 there is a very significant diatribe of 39 sentences. Most of the sentences are short and include a considerable number of accusing questions. There are three examples of the AmpliClRt (not in the questions) in which both parts are in the POT aspect. A fourth AmpliClRt, which speaks of what has already been said and has three PSs, has the three aspects: CMP + CNT + POT (cf. (13)). It may very well be that POT aspect in SMM is especially characteristic of a diatribe.

Much more study is needed to find real relevancy in the distribution of aspect in the AmpliClRt. This would include studying aspect and time units in the larger contexts of the AmpliClRt.

7. SUMMARY. AmpliClRt: (a) a grammatical unit consisting of 2 or 3 simple predicate-subject sequences which fills the nucleus of a clause, hence is a clause root; (b) a phonological unit; (c) uses largely directional motion verbs; (d) contrasts with the multiple verb complexes in a VP; (e) presents the need to further study the relevancy of aspect in its VP.

Data from 72 examples from total corpus			**Data from 36 examples from Text 54**		
Same aspect in each PS:	27	38%	*Same aspect in each PS:*	16	44%
CMP + CMP:	13	18%	CMP + CMP	8	22%
POT + POT:	6	8%	POT + POT	7	19%
CNT + CNT:	4	6%	CNT + CNT	1	3%
Different aspects in each PS:			*Different aspects in each PS:*		
CMP + POT	15	**21%**	CMP + POT	11	**31%**
POT + CMP	1	1%	POT + CMP	0	
CMP + CNT	11	15%	CMP + CNT	4	11%
CNT + CMP	2	3%	CNT + CMP	0	
POT + CNT	0 (one ambiguity)		POT + CNT	0	
CNT + POT	4	6%	CNT + POT	2	6%
CMP *in one or both PSs*	42	**58%**	CMP *in one or both PSs*		**64%**

Table 1. *Combination of aspect in the two PS sequence.*

The AmpliClRt communicates a vivid picture of the speaker's understanding of a complex happening reflected in an infrequently used grammatical structure, linking it directly to a somewhat more simple happening with its more frequently used structure, since the two fill the same slot. Once again I mention one of the very first instances of this structure that claimed my attention. It was what was said to me about the donkey that was hurrying down the trail in front of me:

> ^{3}*Ndaba -tə̄ cuahān -tə̄.*
> jumpPOT -3AML goPOT -3AML
> 'It went bouncing along.'

This is a very much more visually accurate picture of that donkey than a more typical construction could give it.

This construction, from my limited experience, is probably quite unique among languages in general, hence merits some attention. It differs from the serial verb constructions that I have worked with in languages of Nepal and Papua New Guinea. Much more investigation is needed to appreciate in greater detail its structure, functions, distributions, and semantic environments.

---

[1] Some abbreviations used in this paper: 1F = 1st formal, 1I = 1st informal, 12 = you & I, 2F = 2nd formal, 2I = 2nd informal, 3FM = 3rd formal male, 3FF = 3rd formal female, 3I = 3rd informal, 3AML = 3rd animal, ´ = high tone, ¯ = low tone, PL = plural, POT = potential, CMP = completive, CNT = continuative, REP = repetitive, CS = causative, INSIS = insistent, SC = sentence connector, ClRt = clause root, U = undergoer of action, Sc = scope, direction of action, *(asterisk) = morpheme so marked is class B, ⁎(asterisk on line) = CNT morpheme B class which has no form other than triggering a tone shift, - (hyphen) = a border between morphemes within a word, a dot between English words = the meaning of a single Mixtec morpheme.

[2] The number preceding the period is the number of the discourse in my corpus from which the example is taken. The number following the period is the unit within that discourse.

REFERENCES

Alexander, Mary Ruth. 1988. A syntactic sketch of Ocotepec Mixtec. In *Studies in the syntax of Mixtecan languages 1*, ed. by C. Henry Bradley & Barbara E. Hollenbach, 151–304. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington.

Longacre, Robert E. 1966. Trique clause and sentence: A study in contrast, variation, and distribution. *International journal of American linguistics* 32: 242–52.

Pike, Kenneth L. 1948. *Tone languages* (University of Michigan publications in linguistics 4). Ann Arbor: University of Michigan Press.

——— & Evelyn G. Pike. 1982. Grammatical analysis. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington.

——— & ———. 1983. Text and tagmeme. Norwood, New Jersey: Ablex Publishing Corporation.

Pride, Kitty. 1965. *Chatino syntax* (Summer Institute of Linguistics publications in linguistics and related fields 12). Norman, Oklahoma: The Summer Institute of Linguistics and The University of Oklahoma.

* * *

DOES PERCEPTION PRECEDE PRODUCTION? EVIDENCE FROM KOREAN-ENGLISH BILINGUALS

Wendy Baker & Pavel Trofimovich
University of Illinois, Urbana-Champaign

INTRODUCTION. Knowing whether and how speech perception and production are related has important implications for research involving, among many other things, first language acquisition (e.g., Vihman 1991), language impairment (e.g., Tallal & Stark 1976), artificial speech recognition (e.g., Rose et al. 1996), as well as linguistic and psychological theories (e.g, Labov et al. 1991). Such a link is not necessarily apparent, considering that the two skills—one involving motor control and the other auditory processing—may or may not be controlled by different mechanisms (e.g., Allport et al. 1987). However, support for the interdependence of perception and production is more positive than negative (e.g., Fox 1982), although the nature of such interdependence is still unknown and heavily debated (Lindblom 1996, Fowler 1996).

One area of research where the relationship between perception and production is less transparent is adult second-language acquisition, which, in comparison to relatively uniform language development in children, is variable along many dimensions (e.g., Mack 1989, Flege et al. 1999). Among the most salient are such individual differences as amount of second-language experience and age at the time of second-language learning. However, by carefully controlling these individual differences, it is possible to gain insights into speech processing in general, and second-language speech perception and production in particular. Understanding such a relationship is important for both theoretical and pedagogical reasons. Theoretically, knowing whether perception does in fact influence production abilities may illuminate how second languages are acquired (Rochet 1995). For pedagogical reasons, such an understanding may help determine what types of second-language training will be most effective, and how foreign accents can be minimized (Bradlow et al. 1999).

In general, three hypotheses about the relationship between speech perception and production have been advanced. In both first- and second-language studies, the most widely supported hypothesis is that accurate perception is at least one necessary component of accurate production (Flege 1995), which translates into perception abilities usually surpassing, and therefore preceding, production abilities, especially for beginning second-language learners. Even advanced language learners, whose perception and production abilities are nearly asymptotic, perceive some vowels more accurately than they produce them (Flege et al. 1999).

The results of second-language acquisition studies, however, are not conclusive. Other researchers working within the same framework have proposed that production and perception are interdependent and develop simultaneously (Best 1995, Fowler 1996). Assuming that perception entails the ability to detect distal articulatory properties of speech (i.e., tongue movement, vocal tract size, etc.), this hypothesis maintains that perception and production are always aligned, so that perception never surpasses, and therefore never precedes, production and vice versa. The McGurk effect (McGurk & MacDonald 1978), the phenomenon that listeners use both visual and auditory information to recover speech, and the results of imitation studies (Ryalls & Pisoni 1997), which illustrate that children rely on imitation when learning to speak, support the hypothesis that perception and production abilities are interdependent. Speech processing models based on this hypothesis predict that speech perception and production abilities in second-language learners develop in synchrony and depend on how difficult it is to recover slight articulatory differences between first- and second-language sounds (Best 1995). In support of this hypothesis, for example, Flege et al. (1997) demonstrated that, at least for some bilinguals, perception and production abilities are aligned.

In contrast, yet another hypothesis states that accurate production actually precedes accurate perception, so that some speech contrasts are maintained in production before they are actually perceived. Many first-language studies of dialect change, for example, have proposed that changes first occur in production before they are sustained in perception (Labov et al. 1991; see, however, Bowie 2001). Second-language studies often demonstrate that at least some learners are able to produce differences between non-native sounds that they cannot perceive (Flege et al. 1997), such as Japanese-English bilinguals who can produce a contrast between /r/ and /l/ without being able to perceive it (Sheldon & Strange 1982). These findings suggest that there is at best only a weak relationship between perception and production, and that the two develop independently.

Conflicting results such as these suggest that the perception/production relationship may change over time. If perception and production abilities are completely aligned initially, the first hypothesis proposes that, as the learner gains experience, perception precedes production. Conversely, the second hypothesis would predict that both perception and production would be aligned regardless of language experience. Finally, the third hypothesis would predict that production may in fact outstrip perception as a second language is learned. Therefore, observing the *development* of perception and production abilities may be essential in differentiating among these three hypotheses. Thus, the primary objectives of this study were (1) to determine which of the three hypotheses most accurately describes the development of perception and production abilities in second-language learners, and (2) to explore what other factors may influence or explain the perception-production link.

1. RELATIONSHIP BETWEEN PERCEPTION AND PRODUCTION. To explore the development of second-language speech perception and production abilities, in this study Korean-English bilinguals performed perception and production tasks of English vowels. To specifically understand how the relationship between the two skills develops, the results from these experiments were correlated and examined on an indi-

	AOA	LOE	Chron. Age	Use of Korean
Late+3	23.96 (3.69)	3.04 (0.40)	27.28 (3.52)	63.00 (20.03)
Late+10	21.35 (2.82)	9.82 (2.26)	32.08 (2.67)	44.00 (22.21)
Early+10	8.98 (1.18)	11.05 (1.88)	20.33 (1.63)	19.00 (20.25)
K. Mono	29.02 (2.68)	0.23 (0.10)	29.25 (2.69)	74.00 (22.21)
E. Mono			25.60 (4.53)	

Table 1.*Means of age of arrival (AOA), length of experience (LOE), chronological age (Chron. Age), and percentage of use of Korean (Use of Korean). Standard deviations appear in parentheses.*

/i/	/i/	/u/	/ʊ/	/æ/	/ɛ/
beat	bit	boot	book	bat	bet
bead	bid	booed	good	bad	bed
heed	hid	who'd	hood	had	head

Table 2. *Vowel stimuli used in the study.*

vidual basis. In Part 1, the influence of age at the time of learning (age) and amount of exposure (experience), as indicators of linguistic experience, on the perception/production relationship was evaluated to understand the developmental component of speech processing. In Part 2, the data were re-analyzed along factors other than age and amount of exposure to gain better insights into the role of individual differences in the perception/production relationship.

1.1 SUBJECTS. Thirty adult Korean-English bilinguals participated in the study. Bilinguals differed in age of first exposure to English, defined as age at the time of arrival in the United States (early vs. late), and the amount of linguistic experience, defined as length of stay in the United States (3 vs. 10 years). In addition, 10 English monolinguals and 10 Korean monolinguals participated as comparison groups for a total of 50 participants altogether. Korean monolinguals had resided in the United States for an average of 3 months and were beginning learners of English. Based on age of arrival and length of residence, the bilinguals were assigned to three groups, with 10 subjects in each group. Table 1 presents demographic variables for the groups used in the study.

1.2. STIMULI. Most previous studies which explore the relationship between second-language perception and production examine difficult consonant contrasts (Sheldon & Strange 1982, Bradlow *et al* 1999). In order to extend the results of these previous studies to vowel contrasts, we chose three English vowel pairs which are confusable and particularly difficult for Korean-English bilinguals to both perceive and produce (Flege et al. 1997): /i/-/i/, /u/-/ʊ/, and /æ/-/ɛ/ placed in the phonetic environments bVt/k, b/gVd, and hVd (see Table 2).

The stimuli used in experiment 1, the speech perception experiment, were produced by three native English speakers and recorded using a Shure unidimensional head-mounted microphone and Sony tcd-d8 dat recorder. All speakers were English monolinguals from Illinois and had had no or minimal exposure to a foreign language. Recordings were digitized at 22,050 Hz, normalized for peak intensity, and ramped off during the first and last 15 msec. to prevent audible clicks. Recorded stimuli were not edited in any other way and were representative of the variability in fundamental frequency, vowel duration, stop-consonant closure duration, and voice-onset time which are found in natural speech. Prior to the experiment, 10 monolingual speakers of English identified the stimuli with a 98% accuracy in a fixed-choice identification task.

1.3. PROCEDURE. In order to obtain a measure of the bilinguals' perceptual abilities, experiment 1 used a fixed-choice identification task. In this task, subjects heard an English word presented over the headphones and matched it, on the computer screen, with one of six response alternatives labeled orthographically as English words. The stimuli on each trial were played individually, with each new trial presented 1 sec. following the subject's response to a previous trial. No feedback and no practice session were given. However, subjects could listen to each token as many times as they wished. Six test stimuli in hVd context recorded by three speakers were presented in three randomized lists for a total of 54 one-stimulus trials per subject. The perception score (or percent correct in identification) obtained in this experiment was used as a measure of bilinguals' perception abilities.

To explore the relationship between the bilinguals' ability to perceive and produce the same sounds, in experiment 2 the same bilinguals performed a word production task. In this task, bilinguals used the word stimuli from experiment 1 to name black-and-white line drawings. For example, the word *bat* was used to name the drawing depicting a baseball bat, whereas the word *booed* was used to name the drawing depicting sports fans at a stadium. While some drawings did not unambiguously depict the objects to be named, the subjects were familiarized with the intended words during the study phase. During the study phase, bilinguals twice viewed each drawing and repeated the intended word recorded by a female native English speaker. The intended word was presented first in a sentence, which provided the necessary context, and then in citation form. In the test phase, which contained two blocks of 18 randomized trials, subjects attempted to spontaneously name the drawings. When no response was given, the expected name was played over a loudspeaker and the subject repeated it. Overall, each subject produced 18 words for a total of 720 words altogether. Only spontaneous productions from the test phase were used in subsequent analyses.

Tape recordings of the production data were further used to obtain word production ratings. Digitized at 22,050 Hz, normalized for peak intensity, and ramped off during the first and last 15 msec. to prevent audible clicks, the 720 word tokens were randomized and recorded onto a high-quality audio tape, with each stimulus presented with a 4-second interval. To obtain word production ratings, this audio tape was played

	E.Mono. (1)		Early+10 (2)		Late+10 (3)	
	mean	s.d.	mean	s.d.	mean	s.d.
Perc.	94.26	4.57	95.19	5.60	59.81	10.69
Prod.	96.39	1.98	91.89	7.20	64.00	10.65

	Late+3 (4)		K.Mono. (5)		ANOVA	
	mean	s.d.	mean	s.d.	F (4.45)	Tukey HSD
Perc.	62.96	18.52	50.37	13.46	31.30**	1,2 > 3,4,5
Prod.	57.67	14.46	50.39	9.53	46.03**	1,2 > 3 > 5

Table 3. *Percentage of vowels perceived and produced correctly by the subjects in the five groups. 'HSD' indicates the results of a Tukey HSD test (alpha = 0.05; ** $p < 0.001$) comparing the performance of the five groups on perception and production tasks when the group factor reached significance.*

to a group of 10 English monolingual listeners for identification. In this task, the English monolingual listeners were instructed to choose, on an answer sheet, one of four response alternatives. For example, for the word token 'bit', the response alternatives were: 'bit', 'beat', 'bet', or 'bid'. The production ratings (or percent correct in production) obtained in this task were used as a measure of bilinguals' production abilities.

1.4. RESULTS AND DISCUSSION. Calculated across vowels and phonetic environments, perception and production data were submitted to two separate one-way analyses of variance (ANOVAs), using age at the time of learning (AOA) and length of experience (LOE) as the grouping variable. Because the group factor (AOA and LOE) reached significance for percent correct in perception [$F(4,45)=31.30$, $p=0.00$] and production [$F(4,45)=46.03$, $p=0.00$], Tukey HSD *post hoc* tests were used to pursue between-group differences (see Table 3). Interestingly, AOA, not LOE, predicted the bilinguals' perception and production performance, therefore suggesting that only early bilinguals are able to both produce and perceive vowels similarly to English monolinguals. Similarly, both groups of late bilinguals performed identically to Korean monolinguals in perception and production.

To test the claim that perception and production abilities may differ for bilinguals assigned to the same age and experience group (within-group differences), separate *t-tests* were performed within each group on perception and production data. Surprisingly, however, perception and production abilities were identical within each group ($p>0.05$), thus indicating that all groups both perceived and produced sounds with the same accuracy. These results may be taken to indicate that perception does not precede production, and that these two perceptuomotor skills develop simultaneously.

To further explore the relationship among the variables, bilinguals' production and perception scores were correlated with demographic variables. Shown in Table 4 (overleaf), the results attest to a strong relationship among the variables. Perception and production data do strongly correlate positively at the 0.71 level. Furthermore, although both correlations are strong and significant, age at the time of learning

	Production	Perception
Perception	0.71 **	
Age or Arrival	-0.82 **	-0.69 **
Use of Korean	-0.60 **	-0.53 **
Length of Experience	0.57 **	0.41

Table 4. *Correlations between perception, production scores and subject variables (alpha=*0.05*; ** p <* 0.001*).*

correlates more strongly with production than with perception accuracy. This suggests that perception is perhaps less susceptible to the influence of age and experience than is production. Assuming that perception and production abilities, at least in some measure, develop as a function of age, this result may also indirectly indicate that perception does, in fact, precede production. Speculative and perhaps premature, such a conclusion, however, merely confirms that perception and production are related; it does not indicate whether one ability precedes the other.

Indeed, the obtained group data and correlations do not provide accurate insights into how perception and production abilities are related, nor how such a relationship develops as a language is learned. Presumably, our failure to detect oft-cited within-group differences between perception and production in bilinguals (Sheldon & Strange 1982, Flege et al. 1999) was an artifact of the grouping variables used in the analyses—age and amount of experience. Therefore, to specifically address this issue, we performed an individual analysis of the data, paying close attention to intra-subject differences in perception and production accuracy.

In close agreement with perceptually based accounts of the development of bilingual perception and production abilities (e.g., Flege 1995, Rochet 1995), we hypothesized that good production would be associated with good perception, and poor production with poor perception. For intermediate learners, such a hypothesis would entail a possible misalignment of both skills (a characteristic of an intermediate learning stage) if accurate perception is indeed an important factor in the development of accurate production.

Based on the bilinguals' production abilities, instead of age or experience, data re-analysis yielded 3 groups of bilinguals: those with good (85-100%), intermediate (70-85%), and poor (less than 70%) production accuracy. In addition, perception and production scores obtained for Korean and English monolinguals were used for comparison purposes. Perception and production accuracy, calculated across vowels and phonetic environments, was examined in two one-way ANOVAs to determine whether the resulting groups indeed differed in these two skills. Presented in Table 5, results included a significant main effect of group for both perception [$F(4,45)=23.34$, $p= 0.00$] and production [$F(4,45)=97.90$, $p=0.00$], indicating that both skills did indeed differ across the three new groups. However, no differences emerged within each group when individual perception and production performance was compared ($p>0.05$), suggesting that perception and production appeared again to be aligned.

	E.mono. (1) (N=10)	Good (2) (N=8)	Interm.+ (3) (N=3)	Interm.- (3) (N=5)	Poor (4) (N=14)	K.mono. (5) (N=10)	ANOVA	
							F (4.45)	Tukey HSD
Perc.	94.26	95.83	93.83	55.19	61.11	50.37	23.34 **	1,2>3,4,5
	(4.57)	(5.84)	(4.28)	(16.58)	(11.28)	(13.46)		
Prod.	96.39	94.86	78.34	75.78	54.48	50.39	97.90 **	1,2>3>4,5
	(1.98)	(3.25)	(5.30)	(5.86)	(9.13)	(9.53)		

Table 5. *Percentage of vowels perceived and produced correctly by Korean-English bilinguals (individual reanalysis). 'Intermediate +' and 'Intermediate –' indicate the bilinguals with intermediate production skills and respectively high and low perception skills. 'HSD' indicates the results of a Tukey HSD test (alpha = 0.05; ** $p < 0.001$) comparing the performance of the five groups when the group factor reached significance.*

Surprisingly however, the bilinguals whose production accuracy was intermediate (75-80%) displayed a rather large amount of variability in their perception scores given similar production performance. As Figure 1 (overleaf) depicts, in half of these bilinguals, perception accuracy was in fact identical to the monolingual perception norm and surpassed production accuracy. Assuming a perceptual basis for accurate speech production, for these bilinguals production ability may, in fact, eventually align with their perception ability and the monolingual production norm. However, the other half of the bilinguals in this group attained better accuracy in production than perception (e.g., Sheldon & Strange 1982). It may be the case that either production precedes perception for these bilinguals, thus lending support to the third hypothesis entertained earlier, or perhaps factors other than perception underlie production accuracy in these individuals.

One such factor might be the ability to perceive contrasts of a speaker's own productions. There is empirical evidence in first-language acquisition that in order to learn to speak, infants, through babbling and cooing, may attempt to match the acoustics of their own productions with the acoustics of others' speech, thus defining categories for native-language sounds (Vihman 1991, Kuhl & Meltzoff 1996). In fact, auditory self-perception, or the ability to hear one's own speech, is deemed critical for establishing accurate perceptual and articulatory representations in early language development. Similar evidence is marshaled in research on language-impaired children, in whom inaccurate language production is frequently associated with an inability to understand one's own speech, given an overall superior ability to understand others (Panagos & King 1975). Findings such as this strongly implicate self-perception as a necessary and perhaps fundamental element in the native-language production ability.

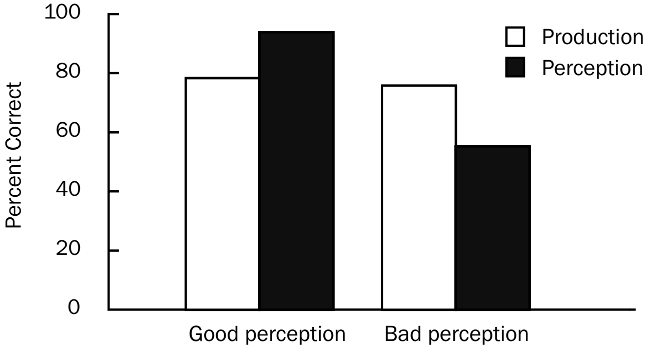


Figure 1. *Percentage of vowels perceived and produced correctly by Korean-English bilinguals with intermediate production accuracy.*

Extended to second-language acquisition, a similar claim may offer a feasible explanation for the somewhat surprising finding of the present study. It may be that auditory self-perception provides the perceptual basis for superior production abilities for those speakers whose production abilities exceed their perception abilities, allowing for the needed mapping of auditory targets onto accurate articulatory targets. More importantly however, from a theoretical perspective, such an explanation would accentuate perceptual bases of accurate production, much in line with most research in the field (e.g., Flege 1995, Rochet 1995).

Beyond doubt, however, auditory self-perception in the absence of the ability to accurately perceive others may not in itself translate into accurate production. Perhaps what self-perception, as part of an auditory-articulatory feedback mechanism, underlies is the ability to *compare* one's own attempts at speech with others' productions. Essential to both first- and second-language learning, self-perception cast in such terms would entail the ability to perceive differences between one's own speech and the auditory targets in others' speech, thus allowing speakers to modify their speech to match these targets.

Because both production and self-perception are contingent upon hearing auditory targets in others' speech, we hypothesize that, for all bilinguals, perception will be better than, or equal to, production and self-perception. By the same token, bilinguals with good production will be equally good at perception and self-perception, because they are able to match their own productions with the auditory targets in others' speech. On the other hand, bilinguals with poor production skills will also be poor at both perception and self-perception because they are unable to match their own productions with the auditory targets in others' speech. However, bilinguals with intermediate production skills will, in large measure, perform better at perception and self-perception than production, because, as hypothesized earlier, the ability to compare one's own speech with auditory targets in others' speech is essential in developing accurate production skills.

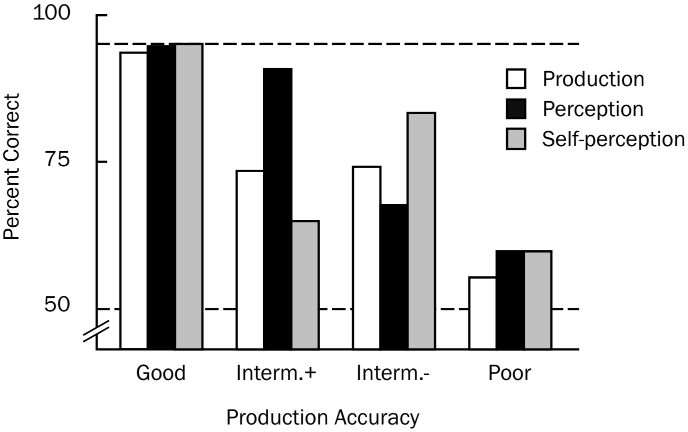


Figure 2. *Production, perception, and self-perception accuracy for Korean-English bilinguals. Dashed lines indicate an average for production and perception performance for English monolinguals (96.39 in production and 94.26 in perception) and Korean monolinguals (50.39 in production and 50.37 in perception).*

2. SELF-PERCEPTION AS A LINK BETWEEN PERCEPTION AND PRODUCTION. Testing these predictions, we asked the same bilinguals to return and perform a word identification task on their own speech to obtain a measure of self-perception. A bilingual's self-perception score was thus defined as the ability to perceive one's own speech identically to the way English monolinguals did in experiment 2. The stimuli and procedure in this experiment were identical to the production rating task performed by English monolinguals in experiment 1. Self-perception scores were available for fourteen bilinguals.

2.1. RESULTS AND DISCUSSION. Chi-square tests comparing perception, self-perception, and production data confirmed our predictions (see Figure 2). More specifically, representing asymptotic performance, the bilinguals whose productions skills are good (N=6) are equally good at perception and self-perception [χ^2=0.80, p>0.05]. Furthermore, the bilinguals whose production skills are poor (N=4) are equally poor at perception and self-perception [χ^2=3.12, p>0.05], performing at chance level and therefore marking the beginning of second-language development. Overall, in both of these stages of second-language learning, perception is completely aligned with production.

The role of self-perception in the developing relationship between perception and production is however most salient in scores of bilinguals with intermediate production skills; i.e., where most of the learning occurs. Of the four bilinguals with intermediate production skills who returned to perform this task, two performed better at perception than production (*Interm.+* group). Indeed, as was predicted, their self-perception and production scores are identical (p>0.05) and intermediate in accuracy. The other two bilinguals in this group performed better at production than perception (*Interm.*- group). Identical to their production accuracy (p>0.05), self-perception accuracy of these bilinguals tended to exceed their perception scores [$\chi^2=3.49$, $p=0.06$]. Implicit in these results is the claim that self-perception, or the ability to distinguish one's own productions from others' speech, is essential in the development of production abilities. Indeed, the bilinguals in the *Interm.+* group belong in the developmental stage in which their self-perception ability is improving, thus enabling them, with more experience, to attain native-like accuracy in a second language. Conversely, the bilinguals whose production surpassed perception accuracy, a somewhat surprising outcome of this study, are good self-perceivers. Being able to perceive their own speech accurately thus enables them to produce sounds accurately. However, these bilinguals may be less likely to attain native-like accuracy in production because, in addition to a good self-perception ability, accurate production should perhaps be grounded in accurate perception.

3. GENERAL DISCUSSION AND CONCLUSIONS. Of the three hypotheses entertained in the introduction, overall this study supports the hypothesis that accurate production is contingent on accurate perception in a second language, and that perception and production abilities may not necessarily be aligned. Indeed, throughout the learning process, perception precedes production. However, we propose that self-perception is important and perhaps even essential in the development of accurate production skills. In fact, self-perception, as both an auditory and articulatory feedback mechanism, may be a necessary link between speech perception and speech production. That is, without self-perception, or the ability to compare one's own attempts at speech with the speech production of others, and regardless of whether such a comparison occurs at an auditory (e.g., Flege 1995) or articulatory (e.g., Best 1995) level, second-language learners may not be able to link auditory perceptual targets with their own articulation efforts. The influence of self-perception may also explain situations where on the surface it appears that production precedes perception. In actuality, as self-perception data demonstrate, accurate production is nevertheless still perceptually based (Newman 1998).

An important implication of this conclusion for speech processing in general and second-language research in particular is that self-perception is essential in learning sounds, enabling speakers to gradually approximate their production ability to their perception ability. Subjecting this hypothesis to more rigorous testing in both first- and second-language research may further illuminate the relationship between perception and production in speech processing.

REFERENCES

Allport, Alan, Donald G. MacKay, Wolfgang Prinz & Eckart Scheerer. 1987. *Language perception and production: Relationships between listening, speaking, reading, and writing.* London: Academic Press.

BEST, CATHERINE T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. by W. Strange, 171–206. Baltimore: York Press.

BOWIE, DAVID. 2001. Perception and production in a series of related mergers. *LACUS forum* 27:297–305.

BRADLOW, ANN R., REIKO AKAHANE-YAMADA, DAVID B. PISONI & YOH'ICHI TOHKURA. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and psychophysics* 61:977–85.

FLEGE, JAMES E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. by W. Strange, 233–77. Baltimore: York Press.

———, OCKE-SCHWEN BOHN & SUNYOUNG JANG. 1997. Effects of experience on nonnative subjects' production and perception of English vowels. *Journal of phonetics* 25:437–70.

———, IAN R. A. MACKAY & DIANE MEADOR. 1999. Native Italian speakers' perception and production of English vowels. *The journal of the Acoustical Society of America* 106:2978–87.

FOWLER, CAROL A. 1996. Listeners do hear sounds, not tongues. *The journal of the Acoustical Society of America* 99:1731–40.

FOX, ROBERT A. 1982. Individual variation in the perception of vowels: Implications for a perception-production link. *Phonetica* 39:1–22.

KUHL, PATRICIA K. & ANDREW N. MELTZOFF. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *The journal of the Acoustical Society of America* 100:2425–38.

LABOV, WILLIAM, MARK KAREN & COREY MILLER. 1991. Near-mergers and the suspension of phonemic contrast. *Language variation and change* 3:33–74.

LINDBLOM, BJÖRN. 1996. Role of articulation in speech perception: Clues from articulation. *The journal of the Acoustical Society of America* 99:1683–1692.

MACK, MOLLY. 1989. Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception and psychophysics* 46:187–200.

MCGURK, HARRY & JOHN MACDONALD. 1978. Auditory-visual coordination in the first year of life. *International journal of behavioral development* 1:229–39.

NEWMAN, ROCHELLE S. 1998. *Individual differences and the link between speech perception and speech production*. Unpublished doctoral dissertation. Buffalo: State University of New York.

PANAGOS, JOHN M. & RELLA R. KING. 1975. Self and mutual speech comprehension by deviant- and normal-speaking children. *Journal of speech and hearing research* 18:653–62.

ROCHET, BERNARD L. 1995. Perception and production of second-language speech sounds by adults. In *Speech perception and linguistic experience: Issues in cross-language research,* ed. by W. Strange, 379–410. Baltimore: York Press.

ROSE, RICHARD C., JUERGEN SCHROETER & M. M. SONDHI. 1996. The potential role of speech production models in automatic speech recognition. *The journal of the Acoustical Society of America* 99:1699–709.

RYALLS, BRIGETTE O. & DAVID B. PISONI. 1997. The effect of talker variability on word recognition in preschool children. *Developmental psychology* 33:441–52.

SHELDON, AMY & WINIFRED STRANGE. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied psycholinguistics* 3:243–61.

TALLAL, PAULA & RACHEL E. STARK. 1976. Relation between speech perception and speech production impairment in children with developmental dysphasia. *Brain and language* 3:305–17.

VIHMAN, MARILYN M. 1991. Ontogeny of phonetic gestures: Speech production. In *Modularity and the motor theory of speech perception,* ed. by I. Mattingly & M. Studdert-Kennedy, 69-84. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

* * *

VOWEL DISCRIMINATION BY SPEAKERS OF ENGLISH, GERMAN, JAPANESE, MANDARIN AND CANTONESE

David C. Bennett
SOAS, London

Zhang Jia-lu, Lü Shi-nan & Hu Xing-hui
Institute of Acoustics, Beijing

Andris Abakuks
Birkbeck College, London

Valérie Hazan
UCL

this paper[1] discusses an extensive cross-linguistic vowel-discrimination experiment and relates the results to various hypotheses concerning speech perception. Section 1 describes the experiment; section 2 presents some of the results; and section 3 gathers together certain points mentioned earlier and relates them briefly to wider issues. The remainder of the introduction is devoted to a brief account of the vowel-discrimination part of the work of Stevens et al. (1969), which will serve as a convenient introduction to our own experiment. Further discussion of their article will, however, be saved until section 3.

Stevens et al. (1969) ran two vowel-discrimination tests, each featuring a 13-step synthetic vowel continuum, with the aim of determining how well native speakers of American English and of Swedish could discriminate pairs of vowels 1 step, 2 steps and 3 steps apart. For one test the continuum involved front unrounded vowels and covered both the English contrast between /i/, /i/ and /ɛ/ and the Swedish contrast between /i/, /e/ and /ä/. The other test involved a continuum taking in the Swedish vowels /i y ʉ/; it thus involved the rounded vowels [y ʉ], which exist in Swedish but not in American English. Figure 1 shows the discrimination functions for the two groups for the 1-step pairs of this test. (The curves for the 2-step and 3-step pairs involved higher percentage scores and were also flatter.) Despite the fact that the Americans had no phonemes corresponding to [y] and [ʉ], they had discrimination peaks in the same place as the Swedes, and the correspondence between the two groups in this test was in fact closer than in the test on unrounded vowels. Stevens et al. (1969:12) interpreted this result as indicating that the discrimination peaks are present universally as part of our hearing mechanism prior to the learning of any particular phonemic system:

> If we extrapolate our results to other vowels and to other languages, the conclusion to be drawn is that, for vowels presented in isolation, the listeners' linguistic experience has essentially no effect upon their ability to discriminate small differences in the vowel formant frequencies.

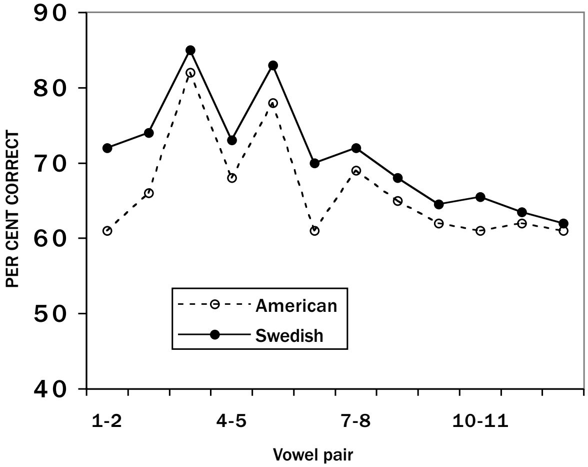


Figure 1. *Discrimination functions for 1-step vowel pairs along a continuum taking in [i y ʉ], as judged by Swedish and American listeners (adapted from Figure 7 of Stevens et al. 1969:11).*

(a)

				⇐ F2
				F1 ⇓
i	y	ɯ	u	
e	ø	ɤ	o	
ɛ	œ	ʌ	ɔ	
a	ɶ	ɒ	ɑ	

(b)

A	B	C	D	E	F	G	H
i		y		ɯ		u	
e	e	ø	ø	ɤ	ɤ	o	o
ɛ	ɛ	œ	œ	ʌ	ʌ	ɔ	ɔ
	a		ɶ		ɒ		ɑ

(c)

I	i	y	ɯ	
J		y	ɯ	u
K	e	ø	ɤ	
L		ø	ɤ	o
M	ɛ	œ	ʌ	
N		œ	ʌ	ɔ
O	a	ɶ	ɒ	
P		ɶ	ɒ	ɑ

Figure 2. *(a) The primary and secondary cardinal vowels, arranged according to F1 and F2; (b) the high-low continua of tests A–H; and (c) the front-back continua of tests I to P.*

In the present experiment we set out to investigate the extent to which such extrapolation is well-founded, by covering all areas of the vowel space—high, low, front, back, rounded and unrounded—and using native speakers of a wider range of languages as subjects, namely English, German, Japanese, Mandarin and Cantonese.

1. THE EXPERIMENT. The synthetic vowels of the experiment include all of Daniel Jones' (1960:31–39) eight 'primary cardinal vowels' [i e ɛ a ɑ ɔ o u] and the corresponding eight 'secondary cardinal vowels' [y ø œ ɶ ɒ ʌ ɤ ɯ] (which have the same tongue position as the primaries but a different lip position). From an acoustic point of view,

(a)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
		intermediate vowels									intermediate vowels								
(e.g.)	[i]									[e]									[ɛ]

(b)	all 19 stimuli paired with themselves, played twice	(38)
	all 18 1-step pairs, in AB and BA order	(36)
	all 17 2-step pairs, in AB and BA order	(34)
	all 16 3-step pairs, played once each, with alternating AB and BA order (1–4, 5–2, 3–6, 7–4…)	(16)

Figure 3. *(a) The stimuli for a given test; and (b) the 124 pairs of vowels of each test.*

on the basis of the frequency of formants 1 and 2, the eight secondary cardinal vowels (CVs) are located between the front series [i e ɛ a] and the back series [ɑ ɔ o u] of the primary CVs, as shown in Figure 2(a).

The experiment comprises sixteen separate tests, each featuring an acoustic continuum taking in three adjacent CVs in either the vertical or the horizontal dimension of Figure 2(a). The eight high-low continua are shown in Figure 2(b) and the eight continua involving front-back (and rounded-unrounded) distinctions are shown in Figure 2(c). It would have been simpler to have organized the synthetic stimuli into only eight tests, one for each of the rows and columns of Figure 2(a). However, we wanted to operate with smaller intervals than Stevens et al. (1969) and synthesized 28 acoustically distinct vowels along each of these dimensions; and because we wanted to present all 3-step, 2-step, 1-step and 0-step (i.e., identical) pairs for judgments, each such test would have taken approximately 30 minutes to run. By using shorter, overlapping continua and doubling the number of tests, we were able to reduce the time taken by each test to around 15 minutes, thereby reducing the burden of concentration on the subjects in any one test.

Figure 3(a) shows how each continuum was divided up. Eight intermediate vowels were synthesized between each of the pairs of adjacent CVs; thus the three CVs of a given test were stimuli numbers 1, 10 and 19. For the CVs themselves the frequencies of the first three formants were derived from analysis of Daniel Jones' recording and for the intermediate vowels the formant frequencies were calculated in such a way that the steps were equidistant logarithmically. The 19 3-formant vowels of each test were synthesized, with a falling intonation (or tone) on a JSRU parallel formant synthesizer at University College London. Each test contains 124 randomized pairs of vowels, made up as indicated in Figure 3(b), and the task is to judge each pair as 'same' or 'different'.

The subjects who took part in the experiment were 5 groups of 16 students, aged in their late teens or early twenties (with the exception of one English girl who was 14 years old). The Mandarin speakers were all natives of Beijing and the Cantonese speakers natives of Hong Kong. The English speakers were all natives of London and the Germans natives of Düsseldorf. Fourteen of the Japanese subjects were natives of Osaka and two came from Kyoto. The tests were played through loudspeakers in a listening room at the Institute of Acoustics in Beijing, in language laboratories in Osaka and London, in a music room in Düsseldorf and in a classroom in Hong Kong.

2. THE RESULTS.

2.1. GENERAL. Over all five groups of subjects, the identical pairs (0-step pairs) received the highest percentage-correct score, averaging out at 80%. Next came the 3-step pairs (73%), then the 2-step pairs (53%), then the 1-step pairs (34%). As also in the case of Stevens et al. (1969), discrimination peaks are more in evidence in the 1-step results. It makes sense, therefore, to concentrate on these (see section 2.2). But first it is of interest to consider the overall performance of each group on the 16 tests.

Figure 4 shows the overall percentage-correct score (across stimulus-pairs of all four types) for each of the 16 tests for each of the five language groups. Analysis of variance (using the statistical package Minitab), carried out earlier on the English, Mandarin and Japanese results and reported in Bennett et al. (1992), revealed *inter alia* that there are highly significant differences between these groups. Moreover, this is true even after allowance is made for differences in the overall level of performance of the groups (which might be due in part to differences in the experimental conditions, such as the rooms in which the tests were played). As regards the ranking of the tests for a particular group, the differences in the percentage scores are often so small that one would expect the ranking to come out differently if the experiment were repeated. Nevertheless, several similarities and differences between the groups are significant. For instance, all groups performed well on tests A and B, covering the range [i e ɛ a], and much less well on tests O and P [a œ ɒ ɑ]. The most striking difference between the groups involves the performance of the English and Mandarin groups on tests I and J, covering the range [i y ɯ u]. For the English subjects, J was the easiest test and I the second easiest. The Mandarin group, on the other hand, performed worst on test I and not much better on test J. A further point of interest is that the English group performed better in general on the front-back tests (I – P) by comparison with the high-low tests, whereas the Mandarin group performed better on the high-low tests than on the front-back tests. In both cases the difference was highly significant statistically. (These points will be taken up again in section 2.4.)

2.2. DISCRIMINATION PEAKS. Figure 5 (overleaf) presents discrimination functions for the 1-step pairs of test A, involving the continuum [i e ɛ], which was one of the tests on which all five groups performed well. It will be seen that all five groups were good at discriminating between stimulus-pairs 3–4, 8–9 and 15–16; and it would be reasonable to suggest that there are discrimination peaks at these three positions along the continuum for all five groups. However, in order to avoid subjectivity in the matter of deciding where there are discriminations peaks, it is desirable to decide on a precise definition of what constitutes a peak. The definition that we chose to employ is conservative rather than liberal, i.e., certain bars in charts such as those in Figure 5 that one might intuitively regard as peaks turn out not to qualify as peaks according to our

English		Mandarin		Japanese		Cantonese		German	
Test	%	Test	%	Test	%	Test	%	Test	%
J: y ɯ u	73.6	B: e ɛ a	57.2	A: i e ɛ	59.5	B: e ɛ a	60.3	B: e ɛ a	67.9
I: i y ɯ	70.4	D: ø œ ɶ	56.6	B: e ɛ a	58.5	A: i e ɛ	60.0	A: i e ɛ	64.4
B: e ɛ a	68.5	*H*: o ɔ ɑ	55.6	I: i y ɯ	57.9	D: ø œ ɶ	58.6	I: i y ɯ	64.2
N: œ ʌ ɔ	68.5	C: y ø œ	55.5	J: y ɯ u	55.2	H: o ɔ ɑ	58.3	J: y ɯ u	64.0
A: i e ɛ	68.0	A: i e ɛ	55.2	H: o ɔ ɑ	55.1	I: i y ɯ	57.6	**G**: u o ɔ	63.1
K: e ø ɤ	67.2	K: e ø ɤ	53.1	C: y ø œ	54.9	G: u o ɔ	55.8	D: ø œ ɶ	61.9
D: ø œ ɶ	65.9	H: o ɔ ɑ	53.0	D: ø œ ɶ	54.3	K: e ø ɤ	55.6	N: œ ʌ ɔ	61.9
M: ɛ œ ʌ	63.4	O: a ɶ ɒ	53.0	K: e ø ɤ	54.2	J: y ɯ u	55.4	**H**: o ɔ ɑ	61.7
H: o ɔ ɑ	63.2	M: ɛ œ ʌ	52.2	G: u o ɔ	53.4	C: y ø œ	54.1	*G*: u o ɔ	61.4
L: ø ɤ o	62.2	N: œ ʌ ɔ	52.0	M: ɛ œ ʌ	52.4	F: ɤ ʌ ɒ	54.0	K: e ø ɤ	61.0
O: a ɶ ɒ	61.8	G: u o ɔ	51.3	N: œ ʌ ɔ	51.5	L: ø ɤ o	53.0	G: u o ɔ	59.2
C: y ø œ	61.4	J: y ɯ u	50.3	E: ɯ ɤ ʌ	51.2	M: ɛ œ ʌ	52.7	L: ø ɤ o	58.4
G: u o ɔ	60.8	F: ɤ ʌ ɒ	50.2	P: ɶ ɒ ɑ	51.2	O: a ɶ ɒ	51.7	C: y ø œ	58.2
F: ɤ ʌ ɒ	57.7	L: ø ɤ o	50.1	F: ɤ ʌ ɒ	50.6	E: ɯ ɤ ʌ	51.6	M: ɛ œ ʌ	57.4
P: ɶ ɒ ɑ	56.5	E: ɯ ɤ ʌ	48.4	L: ø ɤ o	50.3	N: œ ʌ ɔ	51.3	E: ɯ ɤ ʌ	56.1
E: ɯ ɤ ʌ	56.1	P: ɶ ɒ ɑ	47.5	O: a ɶ ɒ	49.9	P: ɶ ɒ ɑ	48.5	F: ɤ ʌ ɒ	54.4
		I: i y ɯ	46.9					O: a ɶ ɒ	50.6
		I: i y ɯ	46.7					P: ɶ ɒ ɑ	48.7

Figure 4. *The 16 tests, A–P, ranked according to the overall percentage-correct score for each of the five groups*[2].

definition[3]. The bars that do constitute peaks according to our definition are shown as dark bars. In the English chart there are two such peaks; the percentage correct score for the pair 3–4 does not differ sufficiently from its two immediate neighbors to count as a peak. In the Mandarin chart there is only one peak, corresponding to the pair 15–16. The other two candidates for the status of peaks in this chart are ruled out not so much by their own percentage values as by the fact that the overall level of performance of this group was relatively high. In the Japanese chart, as in the English chart, two of the three candidates for the status of peaks count as peaks according to the definition and the third, for pair 3–4, does not. In the Cantonese chart, all three candidates count as peaks; and it will be noticed that the overall level of performance is lower than that of the Mandarin group. Finally, the German chart has peaks not only at the same three places as the Cantonese chart but also two additional peaks, for pairs 6–7 and 13–14. What is also notable about the German chart is that, apart from the peaks, the overall level of performance is quite low. It contrasts in this respect with the Mandarin chart.

We may contrast the 1-step discrimination functions for test A (Figure 5, overleaf) with those for test I (Figure 6, overleaf), which covers the range [i y ɯ]. This was a test on which the English group performed well and the Mandarin group performed badly. It will be seen that the English and German subjects had one peak, situated at stimulus-pair 10-11, whereas the Mandarin, Japanese and Cantonese charts exhibit a relatively even performance over the 18 pairs and contain no peaks.

Figure 7 summarizes the number of peaks that were found altogether, according to the definition of peaks given in note 3, and presents the information once for the

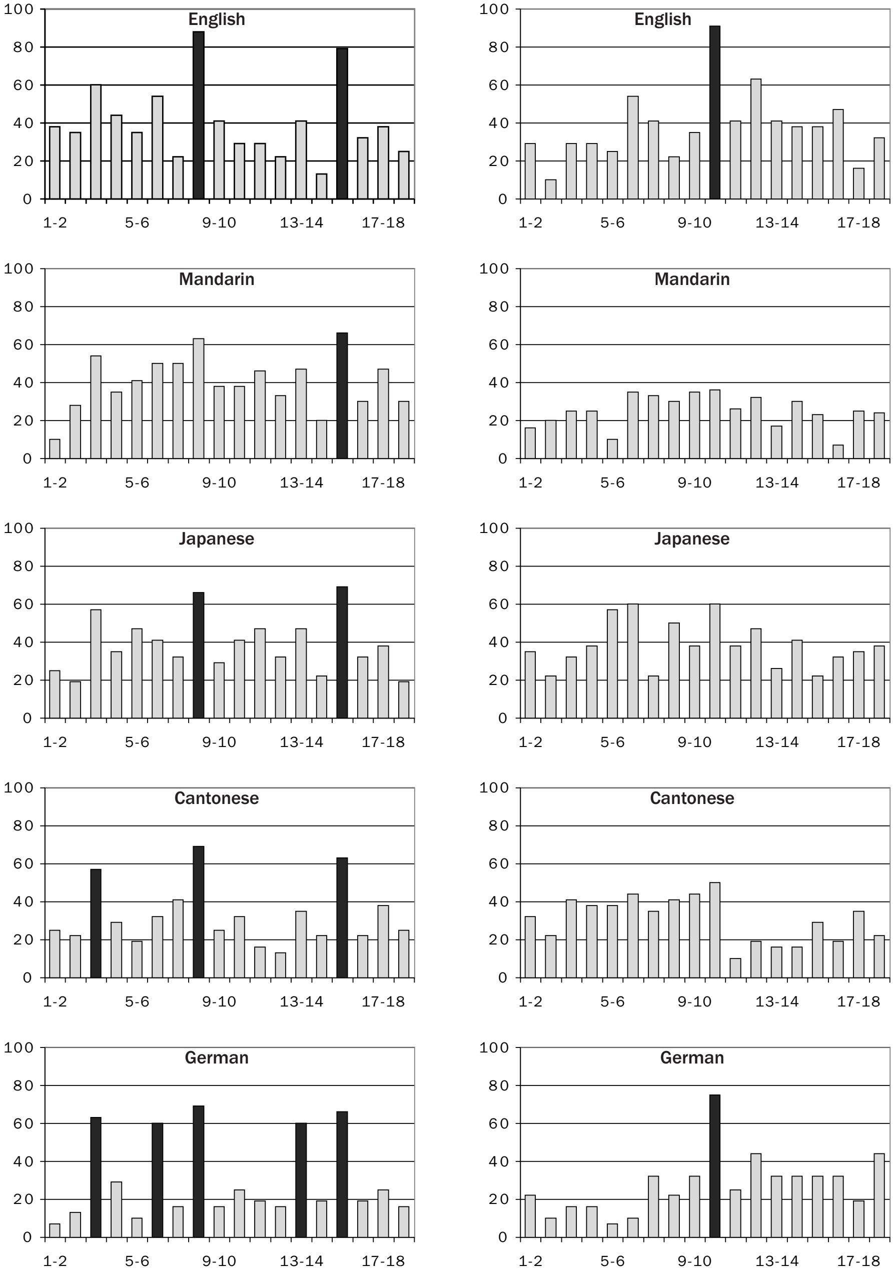


Figure 5 (left column), 1-step discrimination functions for Test A [i e ɛ], and *Figure 6* (right column), 1-step discrimination functions for Test I [i y ɯ]. The x- and y-axes identify vowel pairs and percentage scores respectively.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
13	15	8	7	5	6	10	10	4	5	10	9	8	8	6	4

English	Mandarin	Japanese	Cantonese	German
36	16	14	20	42

Figure 7. *The number of peaks found (a) in the 16 tests and (b) according to language-group.*

individual tests and once for the five language groups. It should be noted that because of the overlap in the range of stimuli covered by the eight pairs of tests A–B, C–D… O–P, some peaks have in effect been counted twice—e.g., the peak for vowel-pair 15–16 in German test A coincides with that for vowel-pair 6–7 in German test B. Certain details of Figure 7 are of some interest, e.g., the fact that the results of the tests that cover the primary CVs [i e ɛ a ɑ ɔ o u], i.e., tests A, B, G and H, exhibit more peaks than are in evidence in the case of the secondary CVs [y ø œ ɶ ɑ ʌ ɤ ɯ], tests C–F. (The front-back tests I–P take in both primary and secondary CVs.) Also of obvious interest is the discrepancy between the number of peaks in evidence in the English and German results and the number of peaks exhibited by the Mandarin, Japanese and Cantonese data. However, limitations of space in the present paper prevent us from exploring such issues further.

2.3. DISCRIMINATION PROFILES. Given that there is a degree of arbitrariness associated with any definition of what constitutes a peak, and given the similarity in profile between the English, Mandarin, Japanese and Cantonese charts in Figure 5 (despite the fact that they differ in the number of peaks present), the possibility suggests itself of comparing the results on the basis of the discrimination profile itself. Spearman's correlation coefficient depends on the ranking of two sets of values rather than the actual values. Bennett et al. (1992) included a table showing Spearman's correlation coefficients for 1-step discrimination functions for all sixteen tests for the language data then available, i.e., the English, Mandarin and Japanese data. The table is reproduced here as Figure 8. For each pair of language-groups, the table indicates the amount of agreement between their responses to each of the tests: the higher a particular test is situated in the diagram, the greater the degree of agreement between the two discrimination profiles in question. The most interesting feature of this table relates to the tests that involved high-low continua, tests A–H. Four of these cover the primary CV continua [i e ɛ], [e ɛ a], [u o ɔ] and [o ɔ ɑ], whereas the other four cover the secondary CV continua [y ø œ], [ø œ ɶ], [ɯ ɤ ʌ] and [ɤ ʌ ɒ]. The primary CV tests are indicated in Figure 8 with a **p** and the secondary CV tests with an **s** on the right of the relevant columns. It will be noticed that there is better agreement between the responses of the three groups of subjects to the primary CV tests than there is in the case of the secondary CV tests. It was obviously of interest to determine whether this relationship would hold also for other pairs of language-groups. This point will be taken up again in section 3.

significance level	English-Japanese		English-Mandarin		Mandarin-Japanese	
	coefficient	test	coefficient	test	coefficient	test
0.01	0.828	B [e ɛ a] **p**			0.843	A [i e ɛ] **p**
	0.813	G [u o ɔ] **p**			0.791	H [o ɔ ɑ] **p**
	0.754	O [a ɶ ɒ]	0.769	B [e ɛ a] **p**	0.736	L [ø ɤ o]
	0.714	H [o ɔ ɑ] **p**	0.735	H [o ɔ ɑ] **p**	0.723	B [e ɛ a] **p**
0.02	0.616	M [ɛ œ ʌ]	0.622	L [ø ɤ o]		
	0.609	A [i e ɛ] **p**	0.59	A [i e ɛ] **p**		
0.05	0.545	L [ø ɤ o]	0.521	N [œ ʌ ɔ]	0.532	G [u o ɔ] **p**
			0.518	K [e ø ɤ]	0.48	I [i y ɯ]
0.1			0.42	G [u o ɔ] **p**	0.394	D [ø œ ɶ] **s**
	0.34	K [e ø ɤ]	0.35	J [y ɯ u]	0.238	N [œ ʌ ɔ]
	0.34	P [ɶ ɒ ɑ]	0.297	D [ø œ ɶ] **s**	0.233	K [e ø ɤ]
	0.336	D [ø œ ɶ] **s**	0.265	E [ɯ ɤ ʌ] **s**	0.216	O [a ɶ ɒ]
	0.314	C [y ø œ] **s**	0.167	M [ɛ œ ʌ]	0.192	E [ɯ ɤ ʌ] **s**
	0.314	N [œ ʌ ɔ]	0.142	O [a ɶ ɒ]	0.14	C [y ø œ] **s**
	0.313	J [y ɯ u]	0.13	F [ɤ ʌ ɒ] **s**	0.07	P [ɶ ɒ ɑ]
	0.31	E [ɯ ɤ ʌ] **s**	0.094	I [i y ɯ]	0.055	J [y ɯ u]
	0.266	F [ɤ ʌ ɒ] **s**	0.076	C [y ø œ] **s**	0.041	M [ɛ œ ʌ]
	0.237	I [i y ɯ]	-0.224	P [ɶ ɒ ɑ]	-0.11	F [ɤ ʌ ɒ] **s**

Figure 8*. Spearman's correlation coefficients for 1-step discrimination functions for all 16 tests for the English, Mandarin and Japanese subjects.*

significance level	correlation coefficient	compared listener-groups
0.01	0.637	English-German
0.02		
0.05	0.48	Mandarin-Japanese
0.1		
	0.356	Japanese-Cantonese
	0.237	English-Japanese
	0.195	Mandarin-Cantonese
	0.094	English-Mandarin
	0.0003	German-Mandarin
	–0.03	German-Japanese
	–0.121	English-Cantonese
	–0.23	German-Cantonese

Figure 9*. Spearman's correlation coefficients for 1-step discrimination functions for test I [i y ɯ] for all pairs of language-groups.*

2.4. CORRELATIONS BETWEEN THE RESULTS AND THE VOWEL SYSTEMS. If an experiment such as the present one were to reveal no statistically significant differences between the various language groups, it would be reasonable to concur with Stevens et al. (1969:12) that 'listeners' linguistic experience has essentially no effect upon their ability to discriminate small differences in… vowel formant frequencies'. Instead, though, the present experiment has revealed various kinds of statistically significant differences between the groups. It is incumbent upon us, therefore, to attempt to explain how speakers may have acquired the differences in question, and the most obvious candidate for an explanation is that they have acquired them in the course of learning the vowel system, or possibly the phonological system more generally, of their native language. However, it is by no means straightforward to establish the relevant correlations between the experimental results and appropriate details of the vowel systems. In the remainder of this section we hope merely to give some idea of the nature of the problem.

Test I [i y ɯ] provides a good example of the kinds of differences that our experiment has revealed. First, the 1-step discrimination functions for this test (see Figure 6) show a single peak in the same position for the German and English subjects and no peaks for the Mandarin, Japanese or Cantonese groups. Moreover, this grouping of German and English *vs*. Mandarin, Japanese and Cantonese is indicated also quite clearly by Spearman's correlation coefficients for this test, as shown in Figure 9. We see, namely, that there is only a 1 in 100 possibility of the degree of correlation between the English and German subjects in this test being due to chance, whereas the much lower degree of correspondence between either the English or the Germans and the other language groups could well be due to chance. Secondly, in terms of the overall performance of the Mandarin and the English subjects, this test was the one on which the Mandarin subjects performed least well, whereas for the English subjects it was the second easiest test (see Figure 4).

The discrimination peak for the English and German subjects in test I involves stimulus-pair 10–11, which would be located roughly in the middle of [y] rather than, say, near the [i]–[y] boundary. Given that German, Mandarin and Cantonese have a phonemic opposition between /i/ and /y/, whereas English and Japanese do not, there is no obvious reason why precisely English and German should have a discrimination peak in this position.

As for the overall difference in performance on test I by the English and Mandarin subjects, one might perhaps initially have predicted the opposite outcome purely on the basis of the listeners' familiarity with the vowel sounds in question: of the sounds [i], [y] and [ɯ], only the first (and a somewhat diphthongal version of it) occurs in the relevant dialect of English, whereas all three can be heard in the Beijing dialect of Mandarin, as in the morphemes *yì* (30+ ways ambiguous in isolation), *yù* (20+ ways ambiguous in isolation) and *sì* (6+ ways ambiguous). However, the vowel [ɯ] in the dialect in question is necessarily preceded by one of the consonantal onsets written in pin-yin romanization as 's', 'c' and 'z'; and pin-yin in effect treats [ɯ] as an allophone of the /i/ phoneme. An alternative prediction, then, would be that phonemic contrasts would have more effect on vowel discrimination than the particular range of allophones that occur. It might seem, namely, *a priori*, that discrimination peaks could be expected to be in evidence at phoneme boundaries but that allophonic differences, which can presumably be ignored by listeners, would have no corresponding peaks.

This last consideration could be involved in the Mandarin subjects' superior performance on high-low tests by comparison with front-back tests. Beijing Mandarin has only one mid-height vowel phoneme, whose realization varies, depending on its context, between [ɛ] in *yè* (e.g., 'night'), [ə] in *wèn* ('ask'), [ʌ] in *è* (e.g., 'hungry') and [ɔ] in *wò* (e.g., 'fertile'). Such widespread allophonic variation could be the reason why the Mandarin subjects were less good at hearing front-back distinctions than high-low distinctions.

Finally, the extra peaks in the German chart in Figure 5 could relate to the presence in German of a four-way phonemic distinction in the [i e ɛ] area of the vowel space: cf. *bieten*, 'to offer', *bitten* 'to request', *beten* 'to pray' and *Betten* 'beds'. However, if diphthongal counterparts are considered, English also has a four-way distinction: *beat*, *bit*, *bait* and *bet*.

3. CONCLUSION. The remainder of our discussion can be conveniently organized around the distinction between what is language-specific and what is universal in relation to vowel discrimination, and speech perception more generally. While Stevens et al. (1969:13) acknowledged that the identification of speech sounds as examples of the phonemic categories of a language reveals language-specific phoneme boundaries, they claimed that vowel discrimination depends on the (universal) auditory mechanism and is unaffected by whatever system of vowels listeners have learned to produce and identify. More recent writers, e.g., Kluender (1994:188–97) and Rosner and Pickering (1994:106–11), redress the balance and acknowledge the importance of linguistic experience also for the discrimination of speech sounds, partly on the basis of the by now quite extensive research on speech sound discrimination by young infants.

The present experiment has provided evidence of both universal and language-specific aspects of vowel discrimination. The similarity between the English, Mandarin, Japanese and Cantonese charts in Figure 5 (the 1-step discrimination functions for test A [i e ɛ]), in terms of both the location of peaks and the overall profile, would seem to suggest a universal mechanism. What lends itself more to a language-specific interpretation in the case of the German chart in Figure 5 is perhaps not so much the presence of the extra peaks that it contains as the sharp contrast between the very prominent peaks and the otherwise rather low percentage scores across the whole range of vowel pairs. Figure 6 also revealed a striking difference between the English and German charts for test I [i y ɯ] and the charts for the other three language-groups, in the form of a prominent discrimination peak for the English and the Germans in the vicinity of the vowel [y], though we were unable to point to any feature of the English and German vowel systems that would explain this shared peak. The shared peak on the part of the English and the Germans, by comparison with the other three groups, in this one test, is in fact part of a general tendency for the English and Germans to exhibit far more peaks than the other groups (see Figure 7).

It is well known that front vowels with spread lips predominate in languages of the world over front vowels with rounded lips; and that back vowels with rounded lips predominate over back vowels with spread lips. Jones's CV system reflects this fact in its choice of the terminology 'primary' and 'secondary'. It has been suggested that the favored association of tongue and lip position is 'consistent with the view that talkers articulate in order to provide acoustic products that are maximally distinctive' (Kluender 1994: 182). Several details of our experiment are consistent with this universal tendency. Each language-group performed somewhat better in tests involving primary CV continua than secondary CV continua (see Figure 4). Secondly, the tests involving primary CVs yielded more discrimination peaks than the tests involving secondary CVs (see Figure 7). Thirdly, Spearman's correlation coefficients revealed greater agreement between pairs of language-groups for the primary CVs than for the secondary CVs. When this test was applied to all further pairings of the five groups, the result was not quite so dramatic as that of Figure 8, where there is no overlapping whatsoever of the primary tests and the secondary tests. Nevertheless, a very strong tendency remains for there to be closer agreement between the groups in the primary tests than the secondary tests.

We feel that we have only just begun to get to grips with the wealth of data available to us. In future work a major challenge will be to attempt to tease apart those aspects of the data that reflect universal tendencies and those that have a language-specific interpretation.

---

[1] We are grateful for the financial support provided at various stages of this project by the following organizations: the Royal Society, the British Academy, the Sino-British Fellowship Trust, the Central Research Fund of the University of London, and the School of Oriental and African Studies. For help in running the experiment in the various locations we are grateful to pupils of Haidian High School, Beijing, and graduate students of the Institute of Acoustics; to Namsun Song, Yoshima Kajiya and students of the Osaka University of Economics and Law; to Ingrid Lüllau and pupils of the Suitbertus-Gymnasium in Düsseldorf; to Cheung Kwan-hin and students of Hong Kong Polytechnic; and to Adrian Bennett, Gregory Bennett and 14 of their friends. We also gratefully acknowledge the help in London of Dick Hayward, Katrina Hayward, Michael Mann and the late Rodney Mellor; and the invaluable computational assistance of Adrian Bennett.

[2] It will be noticed that there are 18, rather than 16, entries for the Mandarin and German groups. There is a different explanation in the two cases. With the Mandarin group it was decided to run tests H and I both at the beginning and at the end of the sequence of tests so that the two sets of results could be compared for consistency. The second run of the two tests is shown in Figure 4 with an italicized *H* or *I*. Since such repetitions of tests added to the length of what was in any case a very long experiment, it was decided not to repeat any tests with the other four groups. On the first trip to Germany, however, test G was run twice inadvertently and test H was not run at all. (These two tests are recorded on opposite sides of the same cassette.) The mistake emerged when the 'H' data were being analysed: the results were quite out of line with the results of all other tests and other groups, e.g., the subjects seemed to have performed better on the 1-step pairs than on either the 2-step or 3-step pairs. When analysed as a second run of test G (for which the random order of the 124 vowel-pairs is different from that of H), the results no longer

appeared anomalous. This second occurrence of test G is entered into Figure 4 with an italicised *G*. Since at this stage no German data were available at all for test H, a second trip was made to the same school in Düsseldorf to run the two tests G and H with a different, but comparable, group of German subjects. The relevant entries in Figure 4 are distinguished by bold-face type.

[3] The definition may be formulated as follows. Given any of the tests A, B… P, a score is recorded at each of the positions 1, 2… 18. Each such score represents the percentage of correct identifications of the vowel difference out of 32 individual trials. The standard error of the difference between any two such scores, s_1 and s_2, is calculated as

$$\sqrt{((s_1 + s_2)(200 - s_1 - s_2))}/8$$

and the test-statistic for testing the significance of the difference $s_1 - s_2$ is

$$z = 8(s_1 - s_2)/\sqrt{((s_1 + s_2)(200 - s_1 - s_2))}$$

A score at a particular position is identified as a peak if it is significantly greater (at the 5% level) than the score at each of the neighboring positions (or significantly greater than the score of its only neighbor if the position is 1 or 18, at the end of the range). A pair of neighboring scores is identified as a combined peak if they do not differ significantly from each other but are each significantly greater than their other neighbors. A triple of neighboring scores is identified as a combined peak if they do not differ significantly from each other but the two outer ones are significantly greater than their other neighbors. Such a procedure provides a systematic way of identifying peaks along the range of positions. However, it should only be regarded as a reasonable guideline. It does not guarantee to find all of what might be regarded intuitively as peaks, nor is it totally free from the risk of picking up spurious peaks caused by random sampling error.

REFERENCES

Bennett, David C., Zhang Jia-lu, Lü Shi-nan, Hu Xing-hui, Valérie Hazan & Andris Abakuks. 1992. Universal and language-specific aspects of vowel perception. In *Proceedings of the 14th International Congress on Acoustics*, vol. 3, ed. by Li Pei-zi, G7–5. Stuttgart: S. Hirzel Verlag.

Jones, Daniel. 1960. *An outline of English phonetics*, 9th edition. Cambridge: Heffer.

Kluender, Keith R. 1994. Speech perception as a tractable problem in cognitive science. In *Handbook of psycholinguistics*, ed. by M.A. Gernsbacher, 173–217. San Diego: Academic Press.

Rosner, Burton S. & J.B. Pickering. 1994. *Vowel perception and production* (Oxford psychology series 23). Oxford: Oxford University Press.

Stevens, Kenneth N., Alvin M. Libermann, Michael Studdert-Kennedy & S.E.G. Öhman. 1969. Crosslanguage study of vowel perception. *Language and speech* 12:1–23.

* * *

PERCEPTION AND PRODUCTION IN A SERIES OF RELATED MERGERS

DAVID BOWIE
Brigham Young University

IN 1972, LABOV, YAEGER, AND STEINER noted the existence of what have been called 'apparent mergers' or 'near-mergers'—cases in which vowel classes are produced differently by members of a speech community but at the same time are perceived by them as being the same[1]. Since then several studies have confirmed that near-mergers do exist (for example Janson & Schulman 1983; Di Paolo & Faber 1990; Labov et al. 1991; Kontra 1993; Faber & Di Paolo 1995; Bowie 2000). In addition, evidence has been found that similar effects can be found in second language acquisition (Sheldon & Strange 1982) and that the concept can be used to explain certain puzzling phenomena in historical linguistics (Nunberg 1980). It should be noted that the evidence listed here is evidence for a rather counterintuitive proposal, as it posits that speakers are able to produce contrasts they are unable to hear.

This paper presents a report of results from fieldwork in Waldorf, a medium-sized town in Southern Maryland, in order to confirm whether mergers in perception actually do precede mergers in production in the linguistic situation there. Waldorf is a community of about 50,000 people located twenty-three miles to the south-southeast of Washington DC; it has historically been a rural community but is in the process of changing into a suburban one. Previous dialectological studies have placed Waldorf either in the Southern dialect region or at the border of the Southern and Midland regions. This town was chosen as the site for the current study in part because previous fieldwork there had shown that there are mergers in progress among some of the pre-lateral non-low back vowels (specifically, the vowels in the words *pole, pull,* and *pool*). The vowels in these words have been found to be merged only in the pre-lateral context in Waldorf; this is unsurprising, as following /l/ is a context that generally promotes merger. This situation presents an opportunity to test the hypothesis that mergers in perception precede mergers in production not just in situations where the merger is universal, but also in cases in which it is a conditioned merger.

1. METHODOLOGY. Fieldwork among Waldorfians was conducted between the spring of 1998 and the fall of 1999. In this study, twenty-nine native Waldorfians underwent commutation tests to determine the progress of the mergers in the pre-lateral non-low back vowels in both perception and production. The subjects took a series of three commutation tests, each comparing a minimal pair of the vowels under investigation to determine the presence or absence of mergers in perception. To accomplish this, each subject was recorded reading three randomly ordered lists of thirty words each, each made up of minimal pairs of the words under study (*pole/pull,*

pseudonym	sex	year of birth	pull-pole	pull-pool	pool-pole
Theona	f	1919	merged	merged	distinct
Elden	m	1921	merged	merged	distinct
Rowan	f	1929	distinct	distinct	distinct
Gerald	m	1941	distinct	merged	distinct
Randall	m	1942	distinct	merged	merged
Elise	f	1946	merged	merged	merged
Raymond	m	1948	merged	merged	distinct
Jeri	f	1951	merged	distinct	distinct
Pippin	m	1951	merged	distinct	distinct
Melina	f	1954	merged	merged	merged
Bo	m	1956	merged	distinct	distinct
Paulie	f	1957	merged	distinct	distinct
Tex	m	1960	merged	merged	distinct
Rosa	f	1963	distinct	merged	distinct
Roy	m	1965	merged	distinct	distinct
Niels	m	1966	merged	merged	distinct
Torren	m	1967	merged	merged	distinct
Blake	f	1969	merged	merged	distinct
Charles	m	1969	merged	merged	merged
Capri	f	1971	merged	merged	distinct
Dayne	m	1973	merged	merged	distinct
Joanne	f	1977	merged	merged	distinct
Helen	f	1978	merged	merged	distinct
Deanna	f	1979	merged	merged	merged
Dawson	m	1980	merged	merged	merged
Gus	m	1982	distinct	merged	distinct
Kelly	f	1986	merged	merged	merged
Cherokee	f	1988	merged	merged	distinct
Thane	m	1988	merged	merged	merged

Table 1. *Status of mergers in perception by subject.*

pull/pool, and *pole/pool*). After this, demographic information was obtained from the subject, and each subject then listened to the recording of their own voice reading the minimal pair lists. The subjects were then asked to mark down which word they believed they heard themselves saying while listening to the tape; a pre-printed form with checkboxes was provided for this purpose.

After the commutation tests were completed, the answers given by the subjects as to what they believed they heard were checked against a master list. If the subject correctly identified all of the words in a particular minimal pair list, then s/he was identified as not having a merger in perception for that pair. On the other hand, if the subject misidentified any of the words on a particular list, s/he was identified as having a merger in perception in pre-lateral environments for the vowel pair in question[2]. In addition, the words rated by the subjects in the commutation tests were put through a linear predictive coding analysis to provide first and second formant values for each utterance to determine the existence of mergers in production using formant values. In this way, the same utterances were tested for evidence of merger in perception and production, and the danger of differences due to style or setting was avoided.

2. RESULTS FOR MERGERS IN PERCEPTION. Looking at the results of the testing for mergers in perception using apparent time reveals that the mergers have proceeded in a particular order in this speech community—first *pull* and *pool* merged, then shortly thereafter *pull* and *pole* merged, and finally the merger of *pole* and *pool* took place. This is shown in Table 1, which gives the results for merger in perception for each of the pairs tested, arranged by the year of birth of the subjects from oldest to youngest. Note that in this table (as in other tables in this paper that use a merged/distinct classification) *merged* means that the items at the top of the column have collapsed together to such an extent that they are rated as merged, while *distinct* means that they are not merged.

Table 1 shows that, while there was some individual variation before (and in some cases after) each merger took hold completely, the most significant difference in the perception of *pull* and *pole* exists between those born in 1963 or earlier and those born in 1965 or later[3]. For *pull* and *pool,* the most significant split occurs between those born in 1957 or earlier and those born in 1960 or later[4], and the merger of *pool* and *pole* shows a significant split between those born in 1978 and earlier and those born in 1979 and later[5].

3. RESULTS FOR MERGERS IN PRODUCTION. Rating the production of mergers is not quite as clear-cut, as production as measured by formant values is an analog rather than a digital system—that is, it's the result of the relevant placement of several produced phonemes rather than a single merged/distinct rating for the entire group. For the purposes of the current paper, however, a system was set up to make rating the subjects' production as binary as possible while still mathematically rigorous. The first and second formant values of each set of produced vowels were subjected to a Student's t test with a null hypothesis that the vowels were merged in production. If the results of this statistical measure gave a value of $p<0.05$, a standard cutoff value for this sort of testing, for *either* formant, the vowels were considered to be distinct in production. A summary of the results for merger in production for all of the individuals listed in Table 1 is given in Table 2 (overleaf).

The most obvious thing to be seen in Table 2 is that there is much less merger in production than there is merger in perception[6]. Also, the splits separating the speech community's adoption of these mergers in production do not precisely match the speech community's adoption of the mergers in production. The most significant difference in production of *pull* and *pole* appears between those born in or before 1967 and those born later[7] (as compared to the perceptual merger, which was adopted slightly earlier, beginning with those born in 1965). The difference in merging *pull* and *pool* in production occurs between those born in 1967 and earlier and those born

pseudonym	sex	year of birth	pull-pole	pull-pool	pool-pole
Theona	f	1919	distinct	merged	distinct
Elden	m	1921	distinct	distinct	distinct
Rowan	f	1929	distinct	distinct	distinct
Gerald	m	1941	distinct	distinct	distinct
Randall	m	1942	distinct	distinct	distinct
Elise	f	1946	merged	distinct	distinct
Raymond	m	1948	distinct	distinct	distinct
Jeri	f	1951	merged	distinct	distinct
Pippin	m	1951	distinct	distinct	distinct
Melina	f	1954	distinct	distinct	distinct
Bo	m	1956	distinct	distinct	distinct
Paulie	f	1957	distinct	distinct	distinct
Tex	m	1960	distinct	distinct	distinct
Rosa	f	1963	distinct	distinct	distinct
Roy	m	1965	distinct	distinct	distinct
Niels	m	1966	distinct	distinct	distinct
Torren	m	1967	merged	distinct	distinct
Blake	f	1969	merged	distinct	distinct
Charles	m	1969	merged	merged	distinct
Capri	f	1971	distinct	distinct	distinct
Dayne	m	1973	distinct	distinct	distinct
Joanne	f	1977	distinct	distinct	distinct
Helen	f	1978	merged	merged	merged
Deanna	f	1979	merged	distinct	distinct
Dawson	m	1980	distinct	distinct	distinct
Gus	m	1982	distinct	distinct	distinct
Kelly	f	1986	distinct	distinct	distinct
Cherokee	f	1988	merged	merged	distinct
Thane	m	1988	merged	distinct	distinct

Table 2. *Status of mergers in production by subject.*

in 1969 and later[8] (later than the parallel merger in perception, which was adopted beginning with those born in 1960). The adoption of the merger in production of *pole* and *pool,* it should be noted, does not lend itself well to statistical testing (and therefore to comparison with the adoption of the merger in perception) because the only person who exhibits the merger is Helen (born 1978).

4. COMPARING THE OCCURRENCE OF MERGERS IN PERCEPTION AND PRODUCTION. A comparison of Table 1 and Table 2 reveals that there is often a mismatch between production and perception among those tested. A table showing where the status of each merger matched or differed is given in Table 3, where a *D* represents a distinction in both production and perception, an *M* indicates a merger in both perception and production, *production* indicates a merger in production but not perception, and *perception* indicates a merger in perception but not production.

pseudonym	sex	year of birth	pull-pole	pull-pool	pool-pole
Theona	f	1919	*perception*	M	D
Elden	m	1921	*perception*	*perception*	D
Rowan	f	1929	D	D	D
Gerald	m	1941	D	*perception*	D
Randall	m	1942	*perception*	*perception*	D
Elise	f	1946	M	*perception*	*perception*
Raymond	m	1948	*perception*	*perception*	D
Jeri	f	1951	M	D	D
Pippin	m	1951	*perception*	D	D
Melina	f	1954	*perception*	*perception*	*perception*
Bo	m	1956	*perception*	D	D
Paulie	f	1957	*perception*	D	D
Tex	m	1960	*perception*	*perception*	D
Rosa	f	1963	D	*perception*	D
Roy	m	1965	*perception*	D	D
Niels	m	1966	perception	*perception*	D
Torren	m	1967	M	*perception*	D
Blake	f	1969	M	*perception*	D
Charles	m	1969	M	M	perception
Capri	f	1971	*perception*	*perception*	D
Dayne	m	1973	*perception*	*perception*	D
Joanne	m	1977	*perception*	*perception*	D
Helen	f	1978	M	M	*production*
Deanna	f	1979	M	*perception*	*perception*
Dawson	m	1980	*perception*	*perception*	*perception*
Gus	m	1982	D	*perception*	D
Kelly	f	1986	*perception*	*perception*	*perception*
Cherokee	f	1988	M	M	D
Thane	m	1988	M	*perception*	*perception*

Table 3. *Mergers in perception and production by subject.*

One thing that can be seen in Table 3 is that, once a merger progresses through both perception and production (as shown with an *M*), it is almost never completely reversed in both perception and production (which would be shown with a *D*). This is to be expected—these are processes of merger, and reversal of merger is an extraordinary occurrence.

Another thing that comes out of Table 3 is that, although the perception and production of the various mergers and distinctions does match in a bare majority of cases, these are nearly matched by the large number of cases in which there is a mismatch. Most notably, however, of the forty-three cases of a mismatch between perception and production, only *one* of them shows a merger in production and not perception (Helen's *pool* and *pole*). This in itself is quite suggestive evidence that merger in perception generally precedes merger in production in this speech community. However, before the data in Table 3 can be taken as absolute evidence for this speech community that near-mergers exist, and particularly that mergers in production consistently follow corresponding mergers in perception in time, a problem in the data has to be dealt with.

5. THE EXCEPTION—MERGER IN PRODUCTION PRECEDING MERGER IN PERCEPTION.

There are four logical possibilities for the sorts of matches and mismatches in the perception and production of mergers that one could possibly find in a single individual: a merger in both perception and production, a distinction in both perception and production, a merger in perception coupled with a distinction in production, and a merger in production coupled with a distinction in perception. Individuals fitting all four of these possibilities were found in, for example, Kontra's 1993 study of near-mergers, and all of these possibilities were also found in this study. This is puzzling, however, because unlike Kontra's study, the way the current study was set up would be expected to eliminate the possibility of a result showing that merger in production precedes that same merger in perception, as the subjects were tested using their own production. Obviously, if a subject's production was merged, it would be impossible to tell them apart—there would necessarily be a merger in perception in that case. Helen, however, was able to tell her apparently merged-in-production *pull* and *pool* apart.

The answer to this puzzle, obviously, is that since in this study the presence or absence of mergers in production is determined solely from first and second formant values, Helen must be differentiating those vowels by means other than first and second formant values. A closer examination shows that this is, in fact, the case—Helen uses an exaggerated rounding distinction, in that she consistently exhibits an extreme rounding of pole (which presumably depresses the third formant) which allows her to tell the difference between it and her less-rounded pool.

This brings up the question of whether there are other such cases, in which individuals use methods to make distinctions that are not reflected by formant values. Upon listening to the production of all of the Waldorf natives, one finds only two other cases in which methods of distinguishing a pair of these vowel classes that do not involve formant differences are used. One of these is Paulie's production of *pull* on the one hand (which she produces with a very short vowel) and *pole* and *pool* on the other (which she produces with a longer vowel). The other case is Cherokee's production of *pool* (which she produces with a broken vowel[9]) versus *pull* and *pole* (which she produces with unbroken vowels). It should be stressed that these are not cases of hypercorrection (with the variation that is found in such cases), but rather that these are consistent parts of these individuals' linguistic systems.

Two items need to be noted regarding this new data. The first is that Paulie's and Cherokee's unexpected production appears to serve no purpose—the values of the first and second formants are quite enough to tell them apart[10]. The second is that it seems rather odd that these three individuals have adopted three such different methods of differentiating these vowels aside from the manipulation of formant values.

These two puzzles can be dealt with by pointing to the same fact—all three of these people (Helen, Paulie, and Cherokee) produce at least some of the vowels they differentiate oddly very close together. It is only in Helen's case that the statistical test of production results in a claim that a pair she differentiates in an unexpected way is actually produced the same according to formant values, but Paulie's production of *pull* and *pool* and Cherokee's production of *pole* and *pool* both overlap physically (though not significantly).

This points to a possibility that will have to remain a conjecture for the moment, but that solves this puzzle quite neatly, which is based on the fact that all three of these cases occur on the cusp of a merger in production (Bowie 1998). It may be that, when the spaces occupied by two vowels start to collapse so that the vowels may merge in production, the individual learner's linguistic system has the option of focusing on secondary differences. This may occur even to the point of changing the features of one or both vowels so that they can remain separate in production while their formant values merge. This conjecture may seem strange when all of the speakers involved are participants in the merger in perception of at least some of the contrasts involved in this, but Nunberg (1980) and Di Paolo (1992) have found that a contrast in production is enough to allow the individual to learn to differentiate a pair of vowel classes. This is the case even for a case of contrast in production without an accompanying a contrast in perception. The fact that all of the methods used in such a way are unique points to the possibility that these were all abortive attempts to change the system, and therefore they are the sorts of things that would only be found sparsely scattered about in the speech community.

6. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH. The most obvious thing to report from this study is that yet another case has been found in which mergers in perception precede mergers in production temporally, thus providing still more evidence that this is a general process by which mergers spread. The data that has been presented, however, shows some items of interest that go beyond that.

I have previously reported on mergers in perception preceding mergers in production (Bowie 1998, 1999). In conducting this research I have found the need for great methodological caution. It is tempting, when comparing mergers in perception and production, to simply compare results for perception with a simple test for merger in production (such as, for example, F1-F2 plots or native speaker judgments). The discussion of the case in this study in which merger in production appeared at first glance to lead merger in perception, as well as the other two cases in which production involved issues aside from first and second formant values, point out the problems inherent in such an approach. That said, though, there is a danger in going to the opposite extreme. As noted earlier, Helen exhibits an extreme rounding of *pole*. Helen was able to make use of the rounded/unrounded contrast for *pole* and *pool* to differentiate them in perception, but she couldn't do this for *pole* and *pull*. The question arises—should studies such as this one consider *pull* and *pool* as being produced differently? Similarly, if they should, at what degree of difference should the cutoff line be? Future work should deal directly with these issues.

Finally, this paper is the latest report to confirm that mergers in perception precede mergers in production in various environments; one could certainly argue that it is a conclusion that is now established. However, this is only one part of the larger issue of matches and mismatches between speech perception and speech production. A vital issue in the perception/production issue is whether this mismatch occurs in situations other than mergers. One has to wonder whether there is a similar perception/production mismatch when phonemic classes split. If the answer is no, then work will have to be done to explain why exactly there is such a difference—but if the answer is yes, linguistics will be a step closer to developing a theory that can account for all types of phonetic change. In addition, finding out whether there is a perception/production mismatch in such spheres as syntactic and lexical change is equally important. Such research would have the effect of helping to determine whether all transmission of linguistic change functions in the same way, or whether the human mind really does deal with different classes of linguistic processes in very different ways.

---

[1] This is a revised and expanded version of a paper presented at the 20th Deseret Language and Linguistics Symposium in Provo, Utah on 6 April 2000. Many thanks to Hikyoung Lee and two anonymous reviewers for comments on this paper.

[2] This is a very strict standard—even just one misidentified word resulted in the subject being labeled as having a merger in production. The rationale, however, is that if there is a robust distinction between two phonemes (as an example, consider the initial consonants in *mat* and *cat*), there would be absolutely no confusion between them.

[3] 78.57% of the older group merges the sounds versus 93.33% of the younger group, verified as significant by a chi-square test to a level of $p<0.0001$.

[4] With 58.33% of the older group merging the sounds and 94.12% of the younger group, verified by a chi-square test to a level of $p<0.001$.

[5] 13.04% of the older group merges the sounds and 66.67% of the younger group, which chi-square testing shows to be significant to a level of $p<0.001$.

[6] This is not just an artifact of the method used to determine merger in production—an analysis that assumes that *any* overlap of vowel spaces constitutes merger in production also results in the finding that merger in production is more widespread than merger in perception (Bowie 1998).

[7] 12.50% merged by the older group and 53.85% in the younger group, with a chi-square test showing this split significant to a level of $p<0.01$.

[8] The younger group merges these 5.88% of the time, the older group 25.00% of the time; chi-square testing declares this significant to a level of $p<0.0001$.

[9] By a 'broken vowel' I mean a vowel broken in two by a reduction in amplitude, so that *pool* seems to be produced somewhat more like [pu.ul] than the unbroken [pul], though without the actual full syllable break shown.

[10] It should be noted that these individuals' ability to tell vowel classes apart using data other than formant values appears limited even in these cases. For example, Cherokee

consistently produces *pool* with a broken vowel and *pole* with an unbroken vowel, but she is unable to access this to distinguish those two words in perception.

REFERENCES

Bowie, David. 1998. Does perception really lead production? Evidence from a series of related mergers in Southern Maryland. Paper presented at NWAV(E) 27, Athens, Georgia.

———. 1999. Observations on the (ir)reversibility of merger. Paper presented at the Shenandoah Language and Linguistics Symposium, Buena Vista, Virginia.

———. 2000. The effect of geographic mobility on the retention of a local dialect. University of Pennsylvania PhD dissertation.

Di Paolo, Marianna. 1992. Hypercorrection in response to the apparent merger of (ɔ) and (a) in Utah English. *Language and communication* 12:267–92.

——— & Alice Faber. 1990. Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language variation and change* 2:155–204.

Faber, Alice & Marianna Di Paolo. 1995. The discriminality of nearly merged sounds. *Language variation and change* 7:35–78.

Janson, Tore & Richard Schulman. 1983. Non-distinctive features and their use. *Journal of linguistics* 19:321–36.

Kontra, Miklós. 1993. The messy phonology of Hungarians in South Bend: A contribution to the study of near-mergers. *Language variation and change* 5:225–31.

Labov, William, Mark Karen & Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language variation and change* 3:33–74.

Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress*. 2 vols. Report on National Science Foundation contract NSF-GS-3287. Distributed by U.S. Regional Survey, Philadelphia, Pennsylvania. Philadelphia: University of Pennsylvania.

Nunberg, Geoffrey. 1980. A falsely reported merger in eighteenth-century English: A study in diachronic variation. In *Locating language in time and space*, ed. by William Labov, 221–50. New York, New York: Academic Press.

Sheldon, Amy & Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese speakers of English: Evidence that speech production can precede speech perception. *Applied psycholinguistics* 3:243–61.

~

PHONEME AND GRAPHEME: HOW PARALLEL CAN THEY BE?

David G. Lockwood
Michigan State University

INTRODUCTION. In one of the classic papers on the linguistic treatment of writing, the British linguist C. E. Bazell (1956) noted that published literature contained two alternative conceptions of the grapheme. In one of these 'the graphic shape of an allograph is dependent on its graphic surroundings' (Pulgram 1951, as quoted by Bazell [359]). In the other view, found in a study of Old English writing by Stockwell and Barritt (1951) 'two graphs are allographs of the same grapheme if they represent the same phoneme, whether or not they stand in relations of complementary distribution' (Bazell 359). So in the first conception <ph> would be the same sequence of two graphemes in *graph* as in *uphill*, and in both cases utterly distinct from the <f> in *fine*. In the Stockwell-Barritt view, on the other hand, <ph> in the first example would be a coallograph with <f> (along with the <gh> of *rough* and *laugh*), and this would be utterly distinct from the other usage, where <p> and <h> would be separate graphemes by virtue of representing different phonemes.

More recently, the German scholar Manfred Kohrt (1986), has given names to these alternative conceptions. The first is termed the ANALOGICAL VIEW, because it makes the grapheme-allograph relation parallel to that between phoneme and allophone. The alternative is called the REFERENTIAL VIEW, because it is based on the phonological reference of the various graphemes.[1]

1. EVALUATION OF VIEWS. In evaluating these views, it must first be said that students of writing need to include both kinds of relations in any model they adopt. Any reasonable overall model, in other words, must have a place to treat both the variant forms of a 'letter', 'character' or other graphemic unit and a way to discuss and systematize as much as possible the referential relations that exist between written and spoken language. Except in the special case of a perfect one-to-one relation between units of writing and spoken language, however, these cannot be treated all at once. They will involve separate sections or components of the overall graphonomy.

That said, the analogical view is the one that makes the most sense, as long as it is incorporated into a model that also allows one to relate writing to spoken language in a systematic way. Two major reasons for preferring this view are as follows:

- INDEPENDENCE. Certainly no respectable linguist would propose that we make our analyses of sound systems dependent on writing. In the same way, it seems undesirable to make our analyses of writing dependent on phonology.

- COMPREHENSIVENESS. If a grapheme is defined as a set of ways of representing a phoneme, it would seem impossible to include under the grapheme writing systems not based on sound, or even those based on syllables rather than segments.[2]

2. REMAINING PROBLEMS. Having decided which fundamental path to follow, however, one must still consider two serious problems, and both involve certain ways in which graphonomy is less than ideally parallel to phonology.

2.1. FIRST PROBLEM: RECOGNIZING COALLOGRAPHS. One problem concerns how we recognize which physically different written shapes are to be assigned to the same grapheme. The referential model suggests an automatic solution to this problem, even though it may have fairly bizarre consequences. In following the analogical model, it is natural to look to parallelisms with phonology. Since the classic presentation of phonemic methodology by Morris Swadesh in 1934, the assignment of sounds as allophones of the same phoneme has been based on three key principles:

1. Phonetic similarity
2. Non-contrastive distribution[3]
3. Pattern congruity

It should be clear that non-contrastive distribution and pattern congruity can readily be applied in graphemic analysis: coallographs must indeed occur in such a way that they do not contrast in a given writing system, and alternative conceptions can be evaluated for their relative contribution to the regularity of patterning, causing one to prefer a solution which contributes more to such regularity over one that contributes less. But what about the analog of phonetic similarity, namely graphic similarity? Is this a reasonable requirement for marks treated as coallographs?

Consideration of familiar facts about alphabetic writing suggests that no requirement of graphic similarity would be reasonable. Consider capital and lower-case characters in the Roman and similar writing systems. Now it is true that capital and lower case letters are not simply coallographs, because in modern writing and printing they are functionally contrastive. Corresponding capital and lower-case forms must be treated as graphemically distinct, but that does not mean that they are simply different graphemes. Rather, following the practice of Gleason (1961) and Herrick (1966), we need to view capitalization as involving a suprasegmental grapheme present in addition to the letter grapheme. So any differences of shape under capitalization can be seen as analogous to allophones of vowel or consonant phonemes under different conditions of accent or tone. Thus, allography would clearly be involved. In the modern uses of the Roman and Greek alphabets (termed Neo-Roman and Neo-Hellenic by Herrick 1974) capitals differ from corresponding lower-case letters in relative size, but often differ in shape as well. Let us consider the situation in the three major modern scripts that use capitalization, the Neo-Hellenic (as for Modern

		Vowel		
		i	u	a
Consonant	p	ᐱ	ᐳ	ᐸ
	t	ᑎ	ᑐ	ᑕ
	v	ᕕ	ᕗ	ᕙ
	r	ᕆ	ᕈ	ᕋ

Table 1*. Representations of some Inuktitut CV syllables in the Evans script. (Based on Nichols 1996)*

Greek), the Neo-Roman (as for English), and the Neo-Cyrillic (as for Russian)[4]. In each of these, there are letters that exhibit no strong resemblances of shape between the two types, alongside others that do.[5] For the Greek alphabet, half of the 24 letters show distinct shapes: Α/α, Β/β, Γ/γ, Δ/δ, Ζ/ζ, Η/η, Μ/μ Ν/ν, Ξ/ξ, Σ/σ, Υ/υ, Ω/ω. The others show identical or closely resemblant shapes: Ε/ε, Θ/θ, Ι/ι, Κ/κ, Λ/λ, Ο/ο, Π/π, Ρ/ρ, Τ/τ, Φ/φ, Χ/χ, Ψ/ψ. In the Roman, the proportion is smaller but still considerable, with the following ten letters having distinct shapes: A/a, B/b, D/d, E/e, G/g, H/h, L/l, N/n, Q/q, R/r. The other sixteen do not have distinct shapes—C/c, F/f, I/i, J/j, K/k, M/m, O/o, P/p, S/s, T/t, U/u, V/v, W/w, X/x, Y/y, Z/z. In the Cyrillic, we find only three of the 32 letters with distinct capital shapes—А/а, Б/б, Е/е—while the others have almost the same shapes, with capitals differing only in size and a few minor details: В/в, Г/г, Д/д, Ж/ж, З/з, И/и, Й/й, К/к, Л/л, М/м, Н/н, О/о, П/п, Р/р, С/с, Т/т, У/у, Ф/ф, Х/х, Ц/ц, Ч/ч, Ш/ш, Щ/щ, Ъ/ъ, Ы/ы, Ь/ь, Э/э, Ю/ю, Я/я.[6]

These examples illustrate that coallographs do not have to show graphic similarity, though obviously they may. In this sense, coallography is more closely parallel to coallomorphy than it is to coallophony: allomorphs of the same morpheme are often phonemically similar, but in line with the usual view, they do not have to be similar at all. Without a counterpart of phonetic similarity, however, there may be problems in deciding how to group graphic shapes as allographs: we can use similarity when it occurs, but we may not always have it available.

As a further illustration, let us consider some uses of the Evans script for some languages spoken in northern Canada. This script was originally devised in 1840 by a missionary named James Evans for the Ojibwe of Ontario and a form of Cree spoken in the vicinity of Norway House, in what is now Manitoba. According to the account in Nichols 1996, forms of this writing have continued to be used for various dialects of Cree and Ojibwe, and the same basic system has been modified for use with some Inuit[7] languages and some sub-Arctic Athabaskan languages as well.

Nichols's account refers to this system as a syllabary, and apparently Evans himself conceived of it in such terms. Its structural organization, however, differs from that of an undoubted syllabary such as Cherokee or Linear B because the shape of a character indicates the initial consonant, while vowel differences are indicated by the orientation assumed by that shape. See Table 1 for some examples, based on the Inuktitut orthography established by the Inuit Cultural Institute. This shows three vowel qualities, which are all that are needed for Inuktitut. For Algonquian languages, four orientations are generally needed.

In the same book that contains this article, Peter Daniels (1996) expanded the traditional typology of writing systems to include some additional types beyond the traditional logography, syllabary, and alphabet. In particular he includes a category called an ABUGIDA, which was intended to cover systems like Ethiopic (from which the name originates) and various Indian writing systems such as Devanagari. While neither Daniels nor Nichols makes this suggestion, it seems reasonable to conclude that the Evans system exemplifies a variety of the abugida rather than a syllabary. The definition of an abugida (ibid 4) states that in such a system 'each character denotes a consonant accompanied by a specific vowel, and the other vowels are denoted by a consistent modification of the consonant symbols'. It is not clear that 'consistent modification' is intended to be broad enough to include the orientational distinctions central to the Evans system. It is at least somewhat different in that there is no obviously unmarked variety, such as the characters for consonant plus short /a/ in Sanskrit Devanagari. It is proposed here that Daniels' otherwise very useful typology[8] be modified so as to allow an abugida to be either DIACRITIC (like the Ethiopic and Devanagari and other Indian systems), or ORIENTATIONAL (like the Evans system)[9].

The general graphemic analysis of an orientational abugida such as is found in those writing systems based on the Evans script should, it would seem, distinguish graphemes of basic shape, which symbolize the consonants, from the orientational factors which distinguish the vowels, so each form associated with a consonant + vowel syllable would be seen as simultaneously embodying a grapheme of shape and another of orientation. Graphemically, this is quite different from a true syllabary, like the Cherokee, whose typical graphemes represent particular consonant + vowel sequences as wholes. It is also different from an alphabet, which has graphemes of different shapes for consonants and vowels. Another factor which must be considered in relation to the Evans script is the symbolization of consonants that stand in syllable-final position, not followed by any vowel. Generally, the device adopted for these is a symbol of reduced size and suprascripted position placed to the right of the character representing the onset and coda. In the forms of the script used for Inuit languages and for the more easterly Algonquian languages, these suprascript consonants share the same shapes as the non-suprascripted consonantal form, and are placed in the orientation that would indicate a following /a/ in a non-suprascripted occurrence. It seems reasonable in such a case to abstract suprascripting as another graphemic property at work here, symbolizing the lack of a following vowel. (This compares functionally to the usage in Sanskrit Devanagari, where an explicit stroke [termed the *virama*] indicates that no vowel is to be pronounced after the consonant.) This would also allow us to consider the orientation associated with /a/ or no vowel to be the unmarked one, suggesting that only the orientations associated with other vowels require additional simultaneous orientational graphemes. An analysis for a representative portion of the Inuktitut version along these lines is presented in Table 2.

			Vowel			
			i	u	a	Ø
Consonant	p	Shape	ᐱ	ᐳ	ᐸ	ᑉ
		Analysis	ᐸ +UP	ᐸ +RT	ᐸ	ᐸ +SS
	t	Shape	ᑎ	ᑐ	ᑕ	ᑦ
		Analysis	ᑕ +UP	ᑕ +RT	ᑕ	ᑕ +SS
	v	Shape	ᕕ	ᕗ	ᕙ	ᕝ
		Analysis	ᕙ +UP	ᕙ +RT	ᕙ	ᕙ +SS
	r	Shape	ᕆ	ᕈ	ᕋ	ᕐ
		Analysis	ᕋ +UP	ᕋ +RT	ᕋ	ᕋ +SS

Table 2. *Some Inuktitut uses of the Evans script with graphemic analysis.*

		Vowel					
		e	i	o	a	Ø (Ojibwe)	Ø (Cree)
Consonant	p	ᐯ	ᐱ	ᐳ	ᐸ	ᑉ	ᑊ
	t	ᑌ	ᑎ	ᑐ	ᑕ	ᑦ	ᐟ
	c	ᒉ	ᒋ	ᒍ	ᒐ	ᒡ	ᐨ
	k	ᑫ	ᑭ	ᑯ	ᑲ	ᒃ	ᐠ
	r	ᕃ	ᕆ	ᕈ	ᕋ	ᕐ	3

Table 3. *Varieties of the Evans script in representing syllable-final consonants in Algonquian languages. (Based on Nichols 1996)*

When we consider the versions of the Evans script in use for more westerly Algonquian languages (generally Cree dialects spoken to the west of James Bay), however, we find a situation that is roughly analogous, but is different in one crucial detail: the shapes found in suprascripted position for a particular consonant are not relatable in any obvious way to those indicating corresponding onsets, as illustrated in Table 3[10]. This could be treated as a matter of allography, with special shapes occurring when accompied by the suprascripting grapheme, just as Roman, Greek, or Cyrillic letters accompanied by a grapheme of capitalization may assume different shapes, sometimes without an obvious similarity to their other shapes. There is, however, a crucial question affecting the acceptability of such an analysis. It would be acceptable if we can consider aspects of the function of a grapheme in relation to phonemes (or other units of a spoken language) in deciding between logically possible graphemic analyses, but not if we are prohibited from doing so, in the way a phonologist surely ought to be prohibited from deciding between alternate phonemic analyses on the basis of any writing system that might be in use for the language.

Name	Denotation of characters	Examples
LOGOSYLLABARY	Individual words (or morphemes) as well as particular syllables	Chinese Sumerian
SYLLABARY	Particular syllables (with no systematic graphic similarity between characters for phonetically similar syllables)	Linear B Cherokee Kana
CONSONANTARY (or ABJAB)	Consonants only	Phonecian Hebrew
ALPHABET	Consonants and vowels	Greek Roman Cyrillic
ABUGIDA:	Consonant accompanied by a specific vowel, with other vowels denoted by a consistent modification of the consonant symbols	Ethiopic Devanagari (Evans?)
FEATURAL SYSTEM	Distinctive features of segments	Han'gul Pitman shorthand

Table 4. Types of writing by units represented according to Daniels 1990, 1995.

It seems better to allow the consideration of the reference of a graph, then, when it helps us to decide a case like this. One alternative would be not to associate the coda-consonant graphs with the others at all, and the other would be to associate them in a totally arbitrary way.

2.2. SECOND PROBLEM: GENERAL APPLICATION OF THE GRAPHEME CONCEPT. An advantage to the general adoption of the analogical rather than the referential view of the grapheme is that it allows the idea of the grapheme to be used with any type of writing system, whether or not it has reference to the phonology of the corresponding spoken language. With this broadening, however, come problems of how to apply the notion in writing systems which relate to spoken languages in various ways. Daniels' typology of these relationships (summarized in Table 4) marks a definite advance over the traditional one. When it is clarified to allow both diacritic and orientational varieties of abugida, it gives six different varieties of phonographic writing, plus the additional category for the Chinese type, which Daniels terms a LOGOSYLLABARY[11].

In virtually any writing system, however, there is a hierarchical arrangement at least roughly parallel to hierarchies seen in phonology and other strata of a spoken language. In the Roman alphabet we have, for instance, written words composed of letters, and letters composed of various kinds of loops, curls, bars, and the extensions termed 'ascenders' and 'descenders' by printers. Accounts of the Chinese writing system (e.g., Coulmas 1989:95–96) tell us that full characters are composed of one or more base characters, and the Chinese tradition gives us a further analysis of the base characters into strokes of just eight basic kinds. The problem is deciding just where in

Name	Denotation of primary graphemes
SYLLABARY	Syllables: usually a CV combination, but also a vowel alone; or a more complex syllable.
CONSONANTARY (ABJAD)	Consonants.
ALPHABET	Consonants and vowels
ABUGIDA: DIACRITIC	Consonant plus a vowel treated as unmarked in the writing system, supplemented by diacritics indicating marked vowels or the absence of a vowel.
ABUGIDA: ORIENTATIONAL	Consonant (or the absence of a prevocalic C), but typically combined with graphemes realized in terms of orientation specifying the vowel accompanying (or its absence).
FEATURAL SYSTEM	Properties of segments, combinable with one another in specified ways to indicate a whole segment, or with secondary graphemes manifested by more abstract properties such a repetition or shading.

Table 5. *Types of phonographic writing as interpreted in terms of primary graphemes.*

such a hierarchy to apply the grapheme concept. Are the graphemes of English written words, letters, or the subparts of letters? Are those of written Chinese full characters, base characters, or strokes?

It is suggested here that this problem can be solved if we consider the relations between written and spoken language to some extent. Specifically, it is proposed that each writing system has a set of primary graphemes, definable on the minimum hierarchical level that bears a systematic relation to the spoken language. Table 5 shows how this applies to the various kinds of phonographic writing, recognizing that there will also be other kinds of graphemes, including suprasegmental and diacritic ones.

For the Chinese type, it appears that the base character would be the most obvious candidate for the status of primary grapheme. Strokes would be clearly excluded due to their lack of relation to the spoken language (any exceptions found being treatable as diacritic graphemes). Of the remaining candidates, the base character seems to be the best choice because it does typically relate to the spoken language, though this relation can be quite complex, sometimes involving a graphic kind of idiomaticity. Some writing systems, such as that of Japanese, are compound in that they combine distinct sets of primary graphemes of different types in one overall system. So the Kanji part of Japanese writing would be subject to essentially the same analysis as Chinese, but the Kana (Katakana/Hiragana) would have primary graphemes of a syllabic type.

3. CONCLUSION. In summary, it is proposed that the proper way to look at writing systems linguistically involves adoption of the analogical view in preference to the referential one, with the proviso that phonology may be considered for two purposes:

(1) classification of graphs as coallographs in the absence of graphic similarity, and (2) determination of the hierarchical level appropriate for the primary graphemes of a written language.

This view may profitably be compared to positions presented in previous LACUS Fora by two experienced specialists in the study of writing, Peter Daniels and Earl Herrick. Daniels (1992) began the discussion by arguing against a structural graphemics, while Herrick (1995a) replied in support of the concept. From a reading of their positions, however, it seems evident that they were arguing in relation to different conceptions. Daniels fundamentally argued against the referential grapheme, while Herrick argued in favor of the analogical view. Both agreed, in other words, on the infeasibility of the referential conception. Herrick presented the view, however, that a structural graphemics based on the analogical conception was indeed feasible. In a reply to this, Daniels emphasized that he did not view writing as a form of language, so he as a linguist was interested almost exclusively in the study of writing in relation to speech. Herrick's conception of a structural grapheme based on the internal relations of writing was therefore of little interest to him.

The view presented in this paper is closer to that of Herrick in seeing a structural concept of the grapheme based on the analogical view as a valuable one. At the same time, it recognizes the importance of studying the relations between written and spoken forms of language, and finds some consideration of these relations of reference to be valuable in solving some otherwise difficult problems of analysis.

While the proposal to depart from a strict analogical view to allow a limited consideration of reference for special purposes does in fact interrupt the strict parallelism with phonology, it does not seem so unreasonable when we consider the asymmetry that exists between phonology and graphonomy. While every spoken language has a phonology by definition, only some have an associated orthography. While the idea of a written language not associated with a spoken language does not seems theoretically unimaginable, it occurs in practice only in unusual cases, such as when we learn to understand an ancient written language we cannot pronounce, or when a deaf person unable to control the spoken form of a language still learns to read and write. In view of this practical asymmetry, the methodological asymmetries suggested here as appropriate do not seem unreasonable. Furthermore, this proposal is not made arbitrarily, but as a solution to genuine analytical problems.

---

[1] Herrick 1995a/b suggests the term 'graphemic grapheme' for a grapheme according to the analogical view vs. 'phonological-fit grapheme' for one conceived according to the referential view. Herrick admits these terms are somewhat ad-hoc, and the first also seems to prejudge any comparison between them, suggesting that graphemes of the sort Herrick favors are 'the genuine article'. Kohrt's terms are therefore preferred in any discussion comparing these views.

[2] In the chapter on Writing and Language in Anttila 1972 the notion of graphemics suggested is explicitly restricted to the study of alphabetic writing (33). This statement can be puzzling to a reader whose acquaintance with the grapheme is based primarily on the two

chapters on writing in Gleason 1961. Classroom use of Anttila's book as the primary text in teaching historical linguistics always required the present writer to comment about this restricted view of the grapheme. It now seems apparent that Anttila had been influenced strongly by the use of grapheme/allograph in the Stockwell and Barritt referential tradition. Indeed the historical linguist would be primarily interested in writing as evidence for the phonology of ancient languages, and so would have a great interest in the referential relations, but this would not seem to justify adoption of that view overall.

[3] Swadesh, of course, spoke of 'complementary distribution'. The use of the amended term reflects refinements made later, specifically in Hockett 1947. The discussion there concerned morphemic rather than phonemic analysis, but the same general principle also came to be applied in phonology.

[4] The use of capitals characterizes most modern alphabets derived from the Greek, though it was not present in the older forms of Greek or Roman. Other scripts and script-variants showing this usage are Armemian, Coptic, some forms of Irish, and Fraktur (the so-called 'black letter' form of Roman common in German through the first half of the 20th Century). See Herrick 1999 for examples of an unsuccessful 1930s proposal to extend capitalization and italicization into Hebrew, intended as a way to modernize that writing system.

[5] Obviously the resemblances are matters of degree. These divisions are based on the author's best judgement.

[6] Cyrillic shows a fairly striking difference from its alphabetic relatives in this regard. This lack of difference makes printed Cyrillic resemble Roman printing that uses small capitals and precludes the use of small capitals for special uses in Cyrillic. There are greater differences, however, when handwritten rather than printed Cyrillic forms are considered.

[7] In 1999 the Canadian government accorded special recognition to this writing system in connection with the establishment of the Inuit territory Nunavut, carved from a portion of the Northwest Territories. Both a special 46-cent stamp and a special $2 coin issued in commemoration of this occasion contained the name of the territory in both Roman and the Evans script.

[8] As Herrick 1995a points out, this typology is based on the fit of the writing system with the spoken language, rather than on internal properties of the writing system itself. This should not disparage its general usefulness, however, in organizing accounts of the written/spoken relations on which it focuses.

[9] Some uses of the Evans system combine certain diacritics with orientational differences to accommodate richer sets of vowel distinctions. A raised dot, for instance, is used in the Inuktitut form to indicate that the vowel involved is long rather than short (Nichols 1996). So this overall system has features of both the orientational and the diacritic subtypes of abugidas. Indeed the introduction of such diacritics would seem to be the only way to devise a referentially adequate orientational abugida for a language with more than 4 distinct vowels. This is relatively easy when there are sets of vowels distinguished by such properties as length, height, frontness, rounding, tongue advancement, or nasalization.

[10] In principle, the shapes the Western Algonquian coda allographs appear to derive from Pitman shorthand, though not all details of usage are identical. For instance, both systems use short dashed lines in different orientations to symbolize stops, but with differences of reference.

[11] This terminology shows the influence of the idea promoted in DeFrancis 1989 treating Chinese writing as a rather complex, morphologically influenced sort of syllabary rather

than a logographic or morphographic system as has been supposed more traditionally. For a contrasting argument, see Sampson 1994.

REFERENCES

Anttila, Raimo. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan

Bazell, C. E. 1956. The grapheme. *Litera* 3:43–46. [Reprinted in *Readings in linguistics II*, ed. by Eric P. Hamp, Fred W. Householder & Robert Austerlitz, 359-61. Chicago: The University of Chicago Press (1966). Page references to reprint.]

Coulmas, Florian. 1989. The writing systems of the world. Oxford: Blackwell.

Daniels, Peter T. 1990. Fundamentals of grammatology. *Journal of the American Oriental Society* 110:727–31.

———. 1992. Is a structural graphemics possible? *LACUS forum* 18: 528–37.

———. 1995. Reply to Herrick. *LACUS forum* 21:425-31.

———. 1996. The study of writing systems, in Daniels & Bright, 3–17.

——— & William Bright, eds. 1996. *The world's writing systems*. Oxford: Oxford University Press.

DeFrancis, John. 1989. *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawai'i Press.

Gleason, H. A. Jr. 1961. *An introduction to descriptive linguistics,* 2nd ed. New York: Holt, Rinehart & Winston.

Herrick, Earl M. 1966. A linguistic description of Roman alphabets. Hartford Studies in Linguistics.

———. 1974. A taxonomy of alphabets and scripts. *Visible language* 8:5–32.

———. 1995a. Of course a structural graphemics is possible! *LACUS forum* 21:413–24.

———. 1995b. Reply to Daniels's reply. *LACUS forum* 21:432-40.

———. 1999. Toward disambiguating the term 'Roman'. *Visible language* 33:102–37.

Hockett, Charles F. 1947. Problems of morphemic analysis. *Language* 23:321–23.

Kohrt, Manfred. 1986. The term 'grapheme' in the history and theory of linguistics. In *New trends in graphemics and orthography*, ed. by Gerhardt Augst, 80–96. Berlin: DeGruyter.

Nichols, John D. 1996. The Cree syllabary, in Daniels & Bright, 599–611.

Pulgram, Ernst. 1951. Phoneme and grapheme: A parallel. *Word* 7:15–20.

Sampson, Geoffrey. 1994. Chinese script and the diversity of writing systems. *Linguistics* 32:117–32.

Stockwell, Robert P. & C. W. Barritt. 1951. Some Old English graphemic-phonemic correspondences. Washington D.C.: Studies in Linguistics Occasional Papers.

Swadesh, Morris. 1934. The phonemic principle. *Language* 10:117–29.

* * *

DERIVING OT CONSTRAINTS FROM RELATIONAL NETWORK THEORY

William J. Sullivan
University of Florida and Uniwersytet Wrocławski

INTRODUCTION. 'Generative' phonology[1], so called because it was intended as a phonological module compatible with the transformational-generative syntax introduced in Chomsky 1957, was first incarnated in Halle 1959. Halle 1959, based on a part of Jakobson, Cherry & Halle 1953, is limited to the production of 'ideal phonetic forms' of the chains of lexical items that make up the output of a *Syntactic Structures* model of syntax. During the 60's generative phonology grew more complex, resulting in the kind of description provided in Chomsky & Halle 1968. By the mid 70's other descriptive models began to be introduced in order to account for phonological phenomena that were as yet unexplained in the generative family of theories (henceforth C-phon)[2]. Among these later models are autosegmental phonology, metrical phonology, and CV phonology. Yet problems remained, many of them[3] artifacts of Chomskyan theory. Some post-1985 C-phon theories have had good work published in them. Underspecification (US) faces a question first asked in Jakobson, Cherry & Halle 1953, i.e. what is the minimal number of features needed to specify the phonemes of a language[4]. Feature Geometry (FG) faces questions also asked there, e.g., what is the principled, structural relation between features? Sullivan 1998 shows that US and FG, which are distinct theor**ies** in the C-phon family, are implicit in and derivable from a well-developed relational network (RN) theory. That is, they are theor**ems** of an RN theory.

An even more recent theory in the C-phon family is Optimality Theory (OT[5]). Again good work has been done in this theory. Milnes and Wiltshire 2000 (henceforth MW) provide an OT description that predicts stress in certain Russian paradigms and Gavrin 1999 provides an OT description of coda position in the Korean syllable. Sullivan (ms.) describes the same Russian problems in a generalized RN description, executed in a logic-based model of cognitive-stratificational theory, and the present study describes the Korean problems attacked in Gavrin 1999. Taken together, the two articles demonstrate that the OT constraints of MW and of Gavrin 1999 are implicit in the RN description. This indicates that OT, too, is a theor**em** of an RN theory.

1. RN DESCRIPTIONS. Cognitive-stratificational theory is not the oldest RN theory but it is the best developed. The networks described are three-dimensional and consist of lines connected at nodes. The nodes may be defined in neurological terms, as in Lamb 1999. However, I am using a coarser model, as my main interest is in the logic of the architecture that accounts for these examples. In this model the nodes are defined

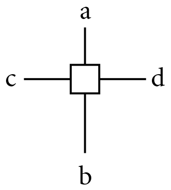


Figure 1. *Logical relations of a phoneme.*

in terms of the logical nature of the relationship (AND, OR, PRECEDENCE). Lines that appear to cross without a node are actually two-dimensional projections of lines that are skew in three-dimensional space. The triangle shows an AND relation, the square bracket an OR. Spacing between lines on the plural side of a node indicates precedence in either time (with AND) or choice (with OR). Here the left-hand line takes precedence in order (with AND) or markedness (with OR). A diamond node is a shorthand for a combination of OR relations that intersect the semiotic (form-function) plane with a structural (tactic) plane, here the phonology. A circle at the end of a line (see figure 2) indicates zero realization; a circle on a line indicates optional occurrence (alternation of what comes below with zero).

I begin the description with the logic of a phoneme. The graphic description is given in Figure 1, which is a maximum potential set of relations for a phoneme. The phoneme itself is represented by the little square box in the center of Figure 1, generally called a diamond. Lines *a-b* are related to points in the structural (phonotactic) part of phonology. Line *a* relates the phoneme to larger (i.e., hierarchically dominant) structures. These structures define its range. Line *b* relates it to smaller (hierarchically subordinate) structures, which define its domain. Conversely, lines *c-d* are in the semiotic (form-function or realizational) part of phonology. Line *c* is related to the set of morphemes that this particular phoneme helps to distinguish from other morphemes, and line *d* is related to its phonetic correlates, initially to the phonemic feature(s). None of the Korean phonemes considered below have more than three of the four relations, but it is easy to exemplify all four, using phonemic stress. With stress, line *a* relates to Pword and line *b* to syllable. The stress phoneme is related to the phonetic correlates of stress (pitch shift, syllable length, loudness, etc.) via line *d*. Finally, line *c* relates it to those morphemes distinctively marked for stress.

2. OT CONSTRAINTS IN KOREAN PHONOLOGY. The OT constraints relevant to Korean phonology and their relative ranking, as described in Gavrin 1999, can be predicted from a generalized RN description of the Korean syllable.[6] The structural and realizational facts are straight-forward. I begin with a summary of syllable structure in Korean.

The Korean syllable permits four canonical forms: V, CV, VC, and CVC. A VCV sequence is syllabified [V CV], where the space indicates the transition between syllables. This shows that onset position of syllable *i+1* takes precedence over coda

Gloss	Maximal form	Phonetic form
'price'	kaps	kap
'soul'	nəks	nək
'1st birthday'	tols	tol
'clothing'	os	ot
'to sit'	anc ta	an t'a
'to tread on'	palp ta	pap t'a
'to recite'	ilp^h ta	ip t'a
'to read'	ilk ta	ik t'a
'to starve'	sulm ta	sum t'a

Table 1. *Coda alternations in Korean.*

position of syllable *i*. Only seven consonants are permitted in coda position: unreleased oral stops /p t k/, nasals /m n ŋ/, and the apical liquid realized as [l][7].

The focus of Gavrin's examples is on what happens with morpheme- and syllable-final consonants. They are realized in onset position, when followed by a morpheme that begins with a vowel, e.g. *-un* 'topic marker'. They are realized in coda position in word-final position or when followed by a morpheme that begins with a consonant, e.g. *-ta* 'infinitive suffix'. In onset position the apical liquid is realized as tapped [r] and other consonants are realized in their phonemic (or underlying) forms. In coda position the consonants have alternate realizations. Some are neutralized with one of the seven permitted consonants, others are realized as zero.

A set of examples taken from Gavrin 1999, representing her native dialect[8], illustrate these alternations. The examples are summarized in Table 1. Recall that all coda stops are unreleased (not indicated in the transcription). C' is a fortis articulation. A space in the 'maximal form' indicates two morphemes, a space in the phonetic form indicates two syllables.

Gavrin's OT description invokes the constraints in (1), ranked as given (a–e).

(1) a. *COMPLEX excludes syllable-internal consonant clusters
  b. PARSE realizes all underlying segments[9]
  c. ALIGN-σ aligns the right[10] edge of the stem with the right edge of the syllable
  d. PERIPHERALITY 'parses' (gives prosodic priority to) labials and velars at the expense of apicodentals
  e. CODA SONORITY parses more sonorous segments at the expense of less sonorous ones

As with the ranked constraints in MW's description of Russian accent, the constraints in (1) produce the correct outputs (or, more properly, weed out the worse output forms).

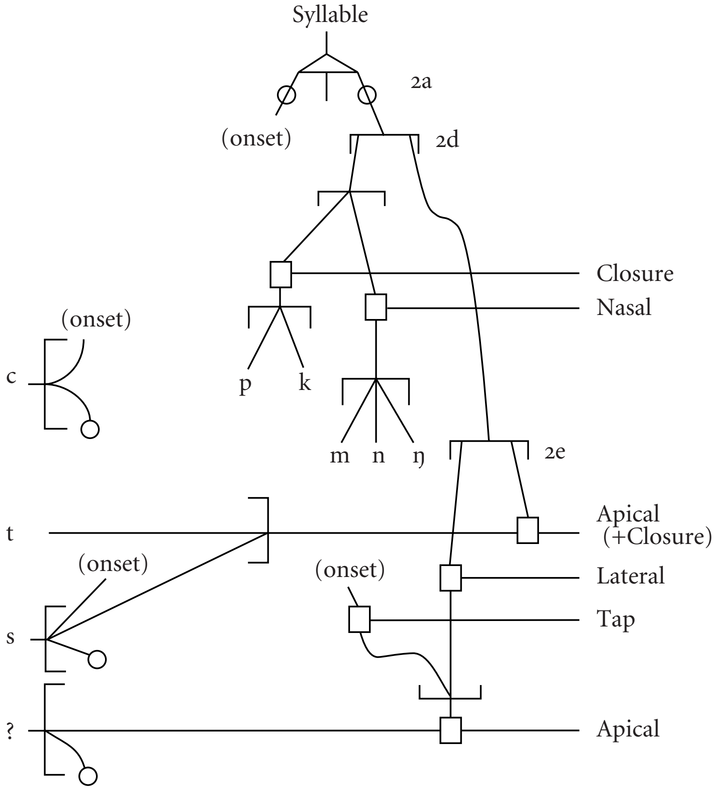


***Figure 2.** Korean coda consonants.*

3. A RN DESCRIPTION OF KOREAN CODAS. The portion of Korean phonology critical to a description of the alternations in Table 1 is given in Figure 2.

Arranged vertically along the left-hand side of the diagram are the relations coming from what might be called the spell-out rules of morphemes. In this case I include three obstruents and the resonant that is realized as either a tapped [r] or an [l]. Reading vertically from the top, I include the relevant portion of the syllable structure, specifically the relations of various phonemes to coda position of the syllable.

Consider the examples in Table 1, focusing on the realizations of underlying *c*, *t*, *s*, and the apical resonant. In onset position *c* is realized as [c], but it is realized as zero in coda position, e.g., [an t'a] ‘to sit’ from *anc ta*. In all cases *t* is realized as [t]. More varied realizations are found for *s*: as [s] in onset position (indicated, not shown), as [t] in [ot] ‘clothing’, from *os*, and as zero in [kap] ‘price’, from *kaps*. The apical resonant has equally varied realizations: as a tapped [r] in onset position (only partially shown), as [l] in coda position, e.g. [tol] ‘first birthday’, from *tols*, and as zero in [ik t’a] ‘to read’, from *ilk ta*. These facts are represented by the horizontal relations in Figure 2, reading from the top to the bottom.

Now return to the structural relations. A generalized statement of the canonical form of the syllable in Korean is [C]V[C] (an optional consonant followed by a vowel followed by an optional consonant).[11] Coda position in Korean accepts only unreleased stops, nasals, and the [l] allophone of the apical resonant. Most liquid-nasal and liquid-stop sequences are realized without the liquid, shown by the ordering at (2d). But if the liquid is followed by an apical stop, whether from a *t*-morpheme or an *s*-morpheme, the sequence is realized without the stop, shown by the ordering at (2e). These are the vertical relations in Figure 2. The bottom-most diamonds are the phonemes that intersect the realizational and tactic relations.

4. DISCUSSION. The RN diagram predicts both the OT constraints and their ranking. *COMPLEX (2a on Figure 2) is inherent in the structure of the AND node labeled 'syllable': both onset and coda positions are optionally related to a single consonant phoneme. Anything more is 'squeezed out' structurally and either is neutralized with something else or is realized as zero.

PARSE is inherent in the realizational relations (2b on Figure 2). To see this, consider the logic of an OR relation like that specifying the realization of the liquid apical in Figure 2. The zero realization is always available. The apical tap [r] is realized in onset position, because the phonotactics supplies the instruction to tap the apex (Tap: from the left-hand diamond related to the upward OR). If no other coda-position consonant (i.e., /p k m n ŋ/) is realized there, the apical lateral [l] is, because the phonotactics supplies the instruction to provide lateral articulation (Lateral: from the right-hand diamond related to the upward OR) in addition to the apical place of articulation. If another coda-position consonant is available, that consonant takes precedence over the apical liquid. Then the other consonant is realized and the apical liquid is realized as zero. Similarly the [c] is realized if it is in syllable onset position. Otherwise it is realized as zero. There is no way to prevent the overt realization of one of these phonemes without adding relations to the diagram, and there is no indication that such additional relations are needed or can be justified. The logic of the network, simple as it is, provides the actual realizations automatically during syllabification.

Now in the OT description, CODA SONORITY parses (realizes) the *n* in *anc ta* 'to sit' and *COMPLEX ensures that the *c* is not realized at all. But the mere occurrence of nasals and not *c* in coda position and the possible zero realization of *c* has the effect of combining the two constraints.

In Figure 2 the ordered OR (2d) ensures that the *m* and not the *l* is realized in [sam] 'life'. This effect is achieved in the OT description by PERIPHERALITY. PERIPHERALITY is a constraint concerning place of articulation, whereas figure 2 shows an ordering based partly on manner of articulation (Nasal) and partly on place of articulation (Apical). A complete and fully generalized description, one that includes onset position, would show the exact extent to which place and manner interact in Korean phonology. Limitations on space and on the author's command of Korean phonology preclude this extension. Still, the effect of PERIPHERALITY also emerges from this partially-generalized RN description.

Excepting only ALIGN-σ, to which I return below, the constraints cited in Gavrin 1999 all emerge from the RN description in Figure 2. The ranking of the constraints also emerges from Figure 2. *COMPLEX and PARSE are hierarchically dominant over all other constraints in Figure 2. In one sense they are logically equivalent, since the output is an intersection of the two sets. But *COMPLEX is a part of the structure of the syllable (2a). The syllable, in turn, provides the necessary conditioning environment for the realizations of the various phonemes. In fact, it governs those realizations. Thus *COMPLEX, as an expression of a part of syllable structure, takes precedence in rank over PARSE. The ordered OR relations (2d and 2e) provide the relative ranking of PERIPHERALITY over CODA SONORITY. They are subordinate to the syllable, and hence to *COMPLEX. Moreover, these relations (and thus PERIPHERALITY and CODA SONORITY) only apply if the phonemes to which the OR nodes are related are activated in the horizontal plane. Thus they are also subordinate to PARSE. This gives the ranking in (2).

(2) COMPLEX >> PARSE >> PERIPHERALITY >> CODA SONORITY

Return now to ALIGN-σ, which aligns the right edge of the stem with the right edge of the syllable. In the examples given in Table 1 (and in Gavrin 1999), this is an accurate requirement. This constraint is not part of any of the structural considerations that went into Figure 2. Properly speaking, it does not actually emerge from Figure 2. Rather, it is an automatic consequence of two facts: the general syllable structure in Figure 2 ([C]V[C]) and the fact that stems in the examples occur either without suffix morphemes or with suffix morphemes that begin with a consonant. It is, in short, an emergent constraint for a restricted set of examples.

At the same time, ALIGN-σ is at least incomplete. If a suffix morpheme begins with a vowel, e.g., *-un* 'topic marker', Figure 2 realizes *kaps un* 'price-TOPIC' as [kap sun]. That is, the stem-final consonant is realized as [s] in onset position of the second syllable. Now the extension of Figure 2 that permits the realization of *s* in onset position provides both realizations, i.e., as [kap] if there is no suffix, and as [kap sun], as remarked above. In sum, the RN description predicts the correct syllabification as well as the alternate realizations of the morphemes simultaneously. If I understand it correctly, an OT description needs syllabification rules in addition to the constraints. Therefore, an RN description incorporates and supersedes ALIGN-σ.

Thus, all of the OT constraints invoked in Gavrin 1999 and their relative ranking emerge from a generalized RN description.

5. REGULAR VARIATION BETWEEN DIALECTS. Gavrin (1999:7–8) cites dialects in which underlying *ilk ta* 'to read' is realized as [il t'a] instead of as [ik t'a]. She accounts for this form by ranking CODA SONORITY above PERIPHERALITY. The RN description of these dialects is simple. The diagram in Figure 2 requires modification of the ordered OR nodes labeled (2d) and (2e). That is, the RN description recognizes that two closely-related dialects may differ phonologically by slight differences in the structure of the syllable[12]. This is a further indication that the ranking between OT constraints

is a direct consequence of phonological structure, as the RN descriptions show. It also explains why the ranking may differ so radically from language to language. The structure of the syllable and Pword differ from language to language, so anything derivable from their structure must also differ from language to language.

This brings us to a consideration of the more general theoretical consequences of OT's findings and RN description.

6. THEORETICAL CONSEQUENCES. In this admittedly restricted but not artificial problem, both the OT constraints and their ranking emerge from a generalized RN description that is focused on more fundamental linguistic considerations, e.g., syllable and morphological structure. This means that Gavrin's description of Korean consonant alternation is actually derivable from the RN description. Sullivan (ms.) shows that this is also true for a radically different problem in an unrelated language (accentology in Russian), as described in MW. I suspect that the same would be true in every case. An actual proof requires the examination of all proposed OT constraints and the relative rankings in the context of particular languages, but we can see some general indications now.

First, OT constraints are of two kinds: universal[13] and specific to a language (or language family). Second, what differs most critically from language to language is the ranking between constraints. As remarked above, OT produces the correct output for the Russian examples in MW and the Korean examples, just as the cognitive-stratificational description does. There is, however, a clear advantage to the RN description in the two cases. For the Russian case, it predicts the correct outcome without explicitly distinguishing between two types of M-stress and establishing a hierarchy between them. More generally, from an O-theoretic point of view, it also shows that the constraint PARSE in Gavrin 1999 and the two constraints STEM and AFFIX in MW are logically identical: a phoneme that alternates with zero.

But this isn't all. Whether OT is a basic theory standing on its own or a sub-theory of C-phon, the existence of the constraints and their ranking are both unexplained 'facts' or theoretical postulates. Conversely, if OT constraints and their relative ranking emerge from an RN description in every case, then OT is actually derivable as a theorem or set of theorems of a fully-developed RN theory. Then the constraints and their ranking follow automatically, as I show in the present case[14]. Moreover, it is an inescapable conclusion that the postulates underlying RN theory are at a much more fundamental level than those underlying OT. Certainly the consequences of these observations relative to the universality of the constraints suggest that the relationship between presumed universals and individual languages requires deeper consideration in the light of relational networks.

In short, it is long past time to turn away from the dead ends that 'mainstream' (Chomskyan) approaches have led to. We need to look seriously at methods of research that use the tools of hard science, neurology among them, and at RN approaches to description. Perhaps then the promise that everyone expected of linguistics a generation ago may at last start to be realized.

[1] It is misnamed. It has always had transformations, initially called Prules, but it never had a generative component, as Sampson 1970 points out.

[2] Nowadays C-phon of the first 15–20 years is often called, rather condescendingly, 'linear' phonology, because more recent models incorporate hierarchical, non-linear features like autosegments (prosodies) and syllables. More recent C-phonologists forget the work of Trubetzkoy, Firth, Pike, Haugen, Hockett, and a host of other phonologists, not the least of whom was Jakobson himself. That is, the C-phonologists merely re-discovered the wheels that had powered phonological research for decades and had been abandoned after 1959.

[3] For example, the prohibition against crossed lines, effectively restricting these models to two-dimensional status.

[4] That is, the minimum number total and the minimum number per phoneme.

[5] Cf. Archangeli & Langendoen 1997. I want to thank Jeong-a Ahn Gavrin of the University of Florida for extended discussions on OT and the chance to read her paper on Korean. I also want to thank Irina Milnes and Caroline Wiltshire of the University of Florida for similar discussions and the opportunity to read preliminary drafts of their article applying OT to Russian stress. Any errors or misinterpretations are my responsibility.

[6] The examples and the OT description are taken from Gavrin 1999. I am dealing with only that portion of her description related to coda 'simplification' in the dialect of Korean that Mrs. Gavrin speaks natively. As a result, the present study does not contain sufficient context to justify its details. Because of space limitations, I present only this portion for the sake of illustration.

[7] When in onset position it is realized as tapped [r]. It has no 'underlying form' because the term 'underlying form' has no meaning in an RN approach. I use *?* as a convenient symbol for it in Figure 2.

[8] I return to the question of dialect variation (Gavrin 1999:7) below.

[9] This constraint subsumes the same functions as stem and affix in MW.

[10] Although this constraint is found in many OT papers, it is, strictly speaking, inaccurate, or at least incomplete, sometimes begging the question of 'resyllabification'. In any event, the actual constraint in this case emerges from the RN description, like all the others. I explain further below.

[11] I am ignoring the fact that the onset of a following syllable takes precedence over the coda. It is easily accounted for in an RN description, but it plays no role in the present study.

[12] Compare the pronunciation of English *ask* in those dialects that neutralize it with *axe*. If this is a general phenomenon in those dialects, the most straight-forward way of accounting for it is to ascribe it to the structure of coda position of the syllable.

[13] It is unlikely that any OT constraint is truly universal, unless it reflects human biology, physiology, or neurology or is based on general logical principles, all of which are non-linguistic. The question of possible universality remains to be determined empirically, but it is irrelevant here.

[14] I am told that C-lin has a number of 'hot' issues (actually unsolved problems) right now, including the source of linear order in texts, the nature of lexical 'features', the source of pronouns, the mechanism of agreement, etc. Answers to all of these problems emerge from an RN approach that needn't even be fully generalized.

REFERENCES

Archangeli, D. & D. T. Langendoen (eds). 1997. *Optimality theory: An overview*. Malden, MA: Blackwell.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

——— & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Gavrin, Jeong-a Ahn. 1999. *The Optimality-Theoretic approach to consonant clusters in the coda in Korean and in Korean loanwords from English.* Unpublished University of Florida MA paper.

Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.

Jakobson, Roman, E. Colin Cherry & Morris Halle. 1953. Toward the logical description of languages in their phonemic aspect. *Language* 29:34–46. [Reprinted in *Selected writings*, vol 1, 449–64. The Hague: Mouton.]

Lamb, Sydney M. 1999. *Pathways of the mind*. Amsterdam: John Benjamins.

Milnes, Irina & Caroline Wiltshire. 2000. Optimal stress patterns in Russian nouns. In *Annual workshop on formal approaches to Slavic linguistics: the Philadelphia meeting 1999*, ed. by Tracy King & Irina Sekerina, 296–313. Ann Arbor: Michigan Slavic Publications.

Sampson, Geoffrey. 1970. On the need for a phonological base. *Language* 46:586–626.

Sullivan, William J. 1998. Underspecification and feature geometry: Theorems of a reticular theory of language. *LACUS forum* 24:53–66.

———. ms. *The predictive power of networks: Describing word accent in Russian*.

SYLLABIFICATION IN PARALLEL IN SPANISH

Caroline R. Wiltshire
University of Florida, Gainesville

SYLLABIFICATION AND SYLLABLE-SENSITIVE alternations in Spanish have long been used as an argument for derivational approaches, in which syllabifications and rules apply in sequence (Harris 1983, 1993, Hualde 1992, Kaisse 1999). For example, a rule that realizes /s/ as [h] is said to apply in some dialects in coda position, if words are syllabified both before and after the rule, as in (1), *desigual* 'unequal' and *las osas* 'the bears'.

(1)	a la Harris (1993)	/des + igual/	/las # osa+s/
	output of first syllabification	(des)(i)(gwal)	(las) (o)(sas)
	s-aspiration in coda	(de[h])(i)(gwal)	(la[h]) (o)(sa[h])
	resyllabification	(de)(hi)(gwal)	(la) (ho) (sah)

Reanalyses in Optimality Theory (Kenstowicz 1996, Colina 1997, Face 1999) avoid resyllabification by using constraints that directly compare output forms, but the use of such constraints is a continuing source of controversy (Hale, Kissock, and Reiss 1998). I propose to analyze Spanish syllabification using an OT approach that relies instead on constraints that preserve the integrity of prosodic constituents. I show how syllabification within and across words can be accomplished in a single step, by the parallel evaluation of such constraints, and handle segmental alternations such as s-aspiration in terms of surface phonotactics, rather than as processes ordered between levels of syllabification. I begin by reviewing the syllabification data in Section 1, followed by my proposal in 2. In 3, I briefly contrast my approach with previous analyses, both rule-based and OT, and conclude with advantages of the approach described here.

1. DATA. Syllabification in Spanish tends to obey two phonological pressures found widely cross-linguistically: syllables prefer onsets over codas, and syllable edges tend to align with word edges. Word-internally, where there is no pressure to align with word edges, the preference for onsets over codas results in the syllabification of single intervocalic consonants in onset position (2)a–b, and the maximization of onset clusters, where possible by the phonotactics (2)c–d.

(2)	a.	/asules/	[.a.su.les.]	*[.as.ul.es.]	'blue'
	b.	/komida/	[.ko.mi.ða.]	*[.kom.i.ða.]	'food'
(2)	c.	/ablar/	[.a.β̲l̲ar.]	*[.aβ̲.lar.]	'talk'

(2) d. /peregrino/ [.pe.ɾe.ɣɾi.no.] *[.pe.ɾeɣ.ɾi.no.] 'pilgrim'

The two pressures conflict, however, when a word-final consonant precedes a word-initial vowel. In that case, a word-final consonant is syllabified into an onset with the following vowel, as in (3).

(3)	a.	/grandes#oxos#asules/	[.gran.de.so.xo.sa.su.les.]	'big blue eyes'
	b.	/asul#oskuro/	[.a.su.los.ku.ro.]	'dark blue'
	c.	/klub#elegante/	[.klu.βe.le.γan.te.]	'elegant club'
	d.	/kalor#insoportable/	[.ka.lo.ɾin.so.poɾ.ta.βle.]	'unbearable heat'

If the following word is consonant-initial, so that it already has an onset, syllabification aligns with the word boundaries, and any word-final consonant is in the coda, (4).

(4)	a.	/klub#lindo/	[.kluβ.lin.do.]	*[klu.βlin.do.]	'beautiful club'
	b.	/čef#loko/	[.čef.lo.ko.]	*[če.flo.ko.]	'crazy chef'
	c.	/benid#rapido/	[.be.nið.ra.pi.ðo.]	*[be.ni.ðra.pi.ðo.]	'come quickly'

This gives the appearance of different rules of syllabification in phrases than word-internally, since a cluster such as /bl/ is treated as a good onset within a word ([.a.βlar.]), but as a coda plus onset in a phrase ([.kluβ.lin.do.]). However, the generalization is that word edges coincide with syllable edges unless a syllable would lack an onset. The same generalization holds across prefix-edges. Unless a syllable would lack an onset (5)a–b, prefix edges coincide with syllable edges (5)c–d.

(5)	a.	/des+igual/	[.de.si.gwal.]	*[.des.i.gwal.]	'unequal'
	b.	/sub+alterno/	[.su.βal.ter.no.]	*[.suβ.a.l.ter.no.]	'subordinate'
	c.	/sub+lu.nar/	[.suβ.lu.nar.]	*[.su.βlu.nar]	'sublunar'
	d.	/ad+risar/	[.að.ri.sar.]	*[.a.ðri.sar.]	'to right (nautical)'

2. ANALYSIS. I first analyze syllabification within simple words and across words. I then extend the analysis to prefixed words, and, finally, show how the structures proposed for syllabification help in analyzing alternations that had been treated as rules serially ordered with resyllabification.

2.1. SYLLABIFICATION WITHIN AND ACROSS WORDS. The analysis of syllabification in phrases is straightforward in an OT account. The pressure for syllables to have onsets is the ONSET constraint (6)a. The pressure for word edges to align with syllable edges is captured by two alignment constraints (6)b and (6)c, parametrized for the left/right edge of the word. Although ultimately low ranking in Spanish, a constraint against codas, (6)d, is included for completeness.

(6) a. ONSET Syllables begin with a consonant.

Input /ablar/	ONSET	NOCODA	ALIGN-R(PW,σ)	ALIGN-L(PW,σ)
☞ [.a.β.lar]	*			
[.aβ.lar]	*	*!		
[.aβl.ar]	**!	**		

***Tableau 1**. Word medial onset maximization*[4].

Input /klub#elegente/	ONSET	NOCODA	ALIGN-R(PW,σ)	ALIGN-L(PW,σ)
☞ [klu.β][e.le.ɣan.te]		*	*	*
[kluβ][.e.le.ɣan.te]	*!	**!		

***Tableau 2**. Phrase-medial misalignment to provide onset.*

Input /klub#lindo/	ONSET	ALIGN-R(PW,σ)	ALIGN-L(PW,σ)	NOCODA
[klu.β][lin.do]		*!	*	*
☞ [kluβ.][lin.do]				**

***Tableau 3**. Phrase-medially, no misalignment to maximze onsets.*

(6) b. ALIGN-L(Wd,σ) The left edge of the word and the syllable coincide.
c. ALIGN-R(Wd,σ) The right edge of the word and the syllable coincide.
d. NOCODA Syllables end with a vowel.

Within a word, ONSET and NOCODA decide the winner, and any ranking of the two results in maximization of the onset cluster (Tableau 1). Across word boundaries, a word-final consonant is syllabified with a word-initial vowel, violating the word-syllable alignment constraints; these are outranked by either ONSET or NOCODA (Tableau 2). As the candidate aligning word and syllable edges has more violations of both ONSET and NOCODA than the candidate that violates word-syllable alignment, we cannot further rank the constraints from this example. However, when a consonant-final word precedes a consonant-initial word, the ranking becomes clear. In these cases, word-syllable alignment is satisfied at the expense of a NOCODA violation (Tableau 3). Thus the ranking that gives us maximal onsets within words and syllabification across words only in cases in which a word would otherwise be onsetless is given in (7). Though the word alignment constraints are outranked, they assert themselves if ONSET is already satisfied.

(7) ONSET >> ALIGN-R, ALIGN-L >> NOCODA

2.2 PREFIXATION. To handle syllabification in prefixed forms, we require an analysis of the structure of prefixed words which is parallel in some sense to that of words in phrases. However, prefixes do not stand alone as words and must be kept structurally distinct from words. Following suggestions in Selkirk (1995) and Peperkamp (1997),

a.	morpheme + suffix	[.a.bl+a+r.]$_{PW}$, [.me.s+eh.]$_{PW}$
b.	prefix + base	[.de.s[i.gual.]$_{PW}$]$_{PW}$, [.sub.[lu.nar.]$_{PW}$]$_{PW}$
c.	words in phrases	[.klub.]$_{PW}$[lin.do]$_{PW}$

***Table 1**. Prosodic structure of words in Spanish.*

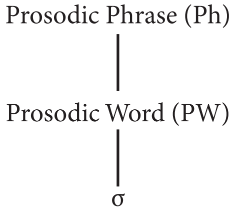


***Figure 1**. Simplified prosodic hierarchy (Selkirk 1984).*

I treat the prosodic word (PW) structure of Spanish as different for prefixed and suffixed words. As shown in Table 1, a suffix coheres to the base within a single PW (row a) with no internal prosodic boundaries. A prefix, however, generates a recursive PW structure, (row b), so that a PW edge intervenes between a prefix and the base. The base is a PW, and the combination of prefix and base is also a PW, but the prefix alone is not. Words in a phrase are distinct PWs, (row c).

Peperkamp (1997) shows that prefixes are not incorporated into the base prosodic word, as suffixes are, based on two rules of Spanish which apply at the left edge of prosodic words: e-epenthesis and r-strengthening. These rules apply regardless of the presence of a prefix before the base, which therefore must begin its own prosodic word[6]. That the prefix is not an entire prosodic word itself is supported by its lack of stress and further motivated by variation in aspiration locations. However, the structure proposed in Table 1(b) violates the Strict Layer Hypothesis (Selkirk 1984), which is based on the Prosodic Hierarchy; a simplified version appears in Figure 1.

The Strict Layer Hypothesis (SLH) proposes that every element lower in the hierarchy is properly dominated by an element one level higher (Selkirk 1984); for use in OT, this has been broken into violable subconstraints (Selkirk 1995, Peperkamp 1997). Selkirk (1995) also argues that the SLH places distinct conditions on lexical words, such as nouns and verbs, vs. non-lexical words, such as function words and prefixes. In alignment terms, such conditions require the edges of lexical words and prosodic words to align, as in Table 2(a) and (b), as well as a constraint that limits recursivity, (c), so that one element cannot dominate another on the same level of the hierarchy.

The recursive structure in Table 1(b) is motivated by the constraint aligning the left edge of lexical words to PWs (Table 2(b)); a base word is a lexical word, while a prefix is not. The ranking of the constraint that requires PWs to end at the end of lexical words prevents prefixes and suffixes from forming PWs themselves (b). To satisfy these constraints, the constraint against recursive PWs, (c), is violated. With (a) and (b) ranked above (c), a recursive word structure is preferred to incorporation of a

a.	ALIGN-L(Lex,PW)	Align the left edge of every lexical word with the left edge of some prosodic word.
b.	ALIGN-L/R(PW,Lex)[8]	Align the left/right edge of every prosodic word with the left/right edge of some lexical word.
c.	NONRECURSIVITY	No C_i dominates C_j, j=i.

***Table** 2.*

Input /sub+lunar/	ALIGN-L(Lex,PW)	ALIGN-L(Lex,PW)	NONRECURSIVITY
[sublunar]$_{PW}$	*!	*(L)	
[sub]$_{PW}$[lunar]$_{PW}$		*(L)*(R)!	
☞ [sub[lunar]$_{PW}$]$_{PW}$		*(L)	*

***Tableau** 4. Phrase-medial misalignment to provide onset.*

Input /mes+es/	ALIGN-L(Lex,PW)	ALIGN-L(Lex,PW)	NONRECURSIVITY
☞ [meses]$_{PW}$		*(R)	
[mes]$_{PW}$[es]$_{PW}$		*(R)*(L)!	
[[mes$_{PW}$]es]$_{PW}$		*(L)	*!

***Tableau** 5. Suffixed word structure.*

Input /sub+ordinado/	ONSET	ALIGN-R(PW, σ)	ALIGN-L(PW, σ)	NOCODA
☞ [.su.b[or.dinado.]]			*	*
[.sub.[or.di.na.do.]]				**

***Tableau** 6. Provide onsets in complex words.*

prefix into the PW with the base (Tableau 4). Suffixes do not interfere with left edge alignment of a lexical word, and are incorporated into the base (Tableau 5).

The structures in Table 1 also allow us a simple explanation of syllabification in the prefixes, using the constraints from 2.1. The constraint forcing the alignment of the left edges of words and syllables, ALIGN-L(PW,σ) , accounts for the failure to syllabify across prefix boundaries as well. As ONSET outranks ALIGN-L(PW,σ), a consonant does syllabify across a prefix (word) boundary if the following PW is vowel-initial, as in Tableau 6. Before consonant-initial words, where ONSET is already satisfied, alignment is preserved at the expense of a NOCODA violation, as in Tableau 7 (overleaf).

In addition to accounting for the parallels of syllabification across word and prefix boundaries, the proposed prosodic word structures enable us to correctly account for the various locations of /s/ aspiration and /n/ velarization in different dialects, to which I now turn.

Input /sub+lu.nar/	ONSET	ALIGN-R(PW, σ)	ALIGN-L(PW, σ)	NOCODA
[.su.β[lu.nar.]]			*	*
☞ [.suβ.[lu.nar]]				**

***Tableau 7**. Don't maximize onsets in complex words.*

2.3. ALTERNATIONS IN PARALLEL WITH SYLLABIFICATION. Spanish syllabification has attracted attention because the allophonics of consonants may depend on word position as well as coda position. This has led to analyses that order syllabification, allophonic rules, and resyllabification. My proposal in 2.2 provides for a surface-based analysis of allophonics such as s-aspiration and n-velarization, and I sketch here a parallel analyses of these two phenomena using constraints at the edges of PWs to account for alternations. In some dialects, both phenomena occur prefix- and word-finally, even if the /s/ or /n/ is in onset positions ([de.hi.gual] 'unequal', [e.ŋa.mo.ra] 'enamors'). First, as in (8), dialects differ on where /s/ is realized as [h], though there is no variation in the realization of /s/ before a consonant; only [h] appears.

(8) Typology of aspirating varieties of Spanish from Kaisse (1998):
All aspirating varieties:

	s → h: morpheme-internal, __C	**stem-final, __C**
a.	*Buenos Aires Argentinian:*	
	s → h: prefix-final, __+C	**word-final, __C**
	[deh[cremada]$_{PW}$]$_{PW}$,	[bičoh]$_{PW}$[raros]$_{PW}$,
	[des[igual]$_{PW}$]$_{PW}$	[bičos]$_{PW}$[ehtraños]$_{PW}$
b.	*Rio Negro Argentinian and Caribbean II:*	
	s → h: prefix-final, __+C	**word-final, __#{C,V}**
	[deh[cremada]$_{PW}$]$_{PW}$,	[bičoh]$_{PW}$[raroh]$_{PW}$,
	[des[igual]$_{PW}$]$_{PW}$	[bičoh]$_{PW}$[ehtrañoh]$_{PW}$
c.	*Caribbean Spanish I (Harris 1993)*	
	s → h: prefix-final, __+{C,V}	**word-final, __#{C,V}**
	[deh[cremada]$_{PW}$]$_{PW}$,	[bičoh]$_{PW}$[raroh]$_{PW}$,
	[deh[igual]$_{PW}$]$_{PW}$	[bičoh]$_{PW}$[ehtrañoh]$_{PW}$

Using the structures developed in 2.2, a suffix coheres to the base within a single PW, with no internal boundaries, so that morpheme-internal and morpheme-final /s/ act the same in all dialects. In a prefixed form, the PW edge which intervenes between a prefix-final /s/ and the base may cause aspiration. Two PW edges follow a word-final /s/, which may also be aspirated. Thus in all the aspirating dialects of Spanish, we need constraints against [s] in preconsonantal position, (9)a, and for some, the constraints against [s] before word edges, (9)b and c, play an active role:

(9) a. *sC No pre-consonantal alveolar fricatives.
b. *s]$_{PW}$: No [s] before the right-edge of PW (i.e., word-finally)

Structure & Inputs	**Candidate Outputs**	***sC**	***h**	FAITHPL	*s|$_{PW}$	*s]$_{PW}$
[morphinternal]$_{PW}$	[dis.co]	*!				
/disc-o/	☞ [dih.co]		*	*		
/visi-o/	☞ [vi.sio]					
	[vi.hio]		*!	*!		
[morph - suf]$_{PW}$	[dies.mo]	*!				
/dies-mo/	☞ [dieh.mo]		*	*		
/mes-es/	☞ [me.seS]		*	*		
	[me.heS]		**!	**!		

Tableau 8*. /s/ →[h] in all aspirating dialects.*

(9) c. *s|$_{PW}$: No [s] before any PW edge (left or right)

When *sC ranks above faithfulness and a constraint barring h, no matter how the other constraints are ranked, morpheme-internal and pre-suffixal /s/'s are realized as [h] in dialects which aspirate. Using the same word and syllable structure for all varieties of Spanish, ranking the constraints in (10)b–c give the differences in dialects found in (8). The three rankings in Tableau 9 (overleaf) have been developed for three types of dialect by Wiltshire (in press).

Thus syllabification remains the same in the different dialects, while the location of aspiration differs. The same holds for /n/ velarization; /n/ is realized in some dialects as [ŋ] in certain positions, including word and prefix finally, as in the Granada dialect (Hualde 1992) in (10).

(10)	PW internal:	a.	/fin+es/	[.fin.eh.]	'ends'
		b.	/enano/	[.e.na.no.]	'dwarf'
	Across words	c.	/sin# fin/	[.siŋ.fiŋ.]	'without end'
		d.	/sin#eso/	[.si.ŋe.so.]	'without that'
		e.	/bien#esta/	[.bje.ŋeh.ta.]	'well-being'
	Across prefixes	f.	/en+amora/	[.e.ŋa.mo.ra.]	'enamors'
		g.	/in+umano/	[.i.ŋu.ma.no.]	'inhuman'

While previous accounts have treated velarization as involving rules interspersed with resyllabification, the same kind of analysis as presented for the aspiration of /s/ can be applied in a parallel approach to these data, as illustrated in Tableau 10 (overleaf).

Again, syllabification is the same across dialects, but the realization of consonantal allophones can differ based on their position in word structure, using a parallel analysis in a single step.

Buenos Aires Argentinian						
Structure & Inputs	**Candidate Outputs**	*sC	*s]$_{PW}$	*h	FAITHPL	*s|$_{PW}$
[pre[morphl]$_{PW}$]$_{PW}$	☞ [de.s[igual]]				*	
/des+igual/	[de.h[igual]]		*!	*!		
[word]$_{PW}$[word]$_{PW}$	☞ [bičo.s][ehtraños]		*	*	*	*
/bičos/ /estraños/	[bičo.h][ehtraños]		**!			
Rio Negro and Caribbean II						
Structure & Inputs	**Candidate Outputs**	*sC	*s]$_{PW}$	*h	FAITHPL	*s|$_{PW}$
[pre[morphl]$_{PW}$]$_{PW}$	☞ [de.s[igual]]					*
/des+igual/	[de.h[igual]]			*!	*!	
[word]$_{PW}$[word]$_{PW}$	[bičo.s][ehtraños]		*!	**	**	*
/bičos/ /estraños/	☞ [bičo.h][ehtraños]			***	***	
Caribbean I						
Structure & Inputs	**Candidate Outputs**	*sC	*s]$_{PW}$	*h	FAITHPL	*s|$_{PW}$
[pre[morphl]$_{PW}$]$_{PW}$	[de.s[igual]]			*!		
/des+igual/	☞ [de.h[igual]]				*	*
[word]$_{PW}$[word]$_{PW}$	[bičo.s][ehtraños]		*!	*!	**	**
/bičos/ /estraños/	☞ [bičo.h][ehtraños]				***	***

Tableau 9*. Variation in /s/ →[h] before vowels.*

Structure & Inputs	**Candidate Outputs**	*n]$_{PW}$	*s|$_{PW}$	*nC	*ŋ	FAITHPL
[morphinternal]$_{PW}$	☞ [.fin.eh.]					
/fin+es/ ‘ends’	[.fi.ŋeh]		*	*		
[pre[morph]$_{PW}$]$_{PW}$	[.e.n[a.mo.ra.]]	*!				
/en+amora/ ‘enamors’	☞ [.e.ŋ[a.mo.ra.]]				*	*
[word]$_{PW}$[word]$_{PW}$	[.si.n[e.so.]]	*!	*!			
/sin/ /eso/ ‘without that’	☞ [.si.ŋ[e.so.]]				*	*

Tableau 10*. /n/ →[ŋ] at prosodic boundaries.*

3. COMPARISON AND CONCLUSIONS. Previous attempts within rule-based approaches and OT have proven to be more complex or less adequate in coverage. Rule-based approaches divide syllabification into stages, building the nucleus, simple onset, complex onset, and coda; these apply in order word-internally (Hualde 1992, Harris 1993). After words are put into phrases, the simple onset rule reapplies, but the complex onset rule does not. As Colina (1997) notes, this is an *ad hoc* restriction, unrelated

to the motivation of the ONSET constraint. Furthermore, for aspiration and velarization, rule-based approaches must order syllabification both before and after aspiration (Kaisse 1999). Though this simplifies the environment for the rule of aspiration, such accounts suffer from unmotivated complexity, adding ordering which results in abstract syllabifications, and adding opacity, since surface [h]'s appear in positions which would not trigger the rule (onset, not coda).

In an OT approach, rather than language particular restrictions on the application of the complex onset rule, we call upon cross-linguistically motivated constraints to distinguish word-internal from cross-word syllabification, and we syllabify only once. In my account, the aspiration and velarization contexts are directly related to the surface forms. Previous accounts in OT (Kenstowicz 1996, Colina 1997, Face 1999) can also claim some of the same advantages: they avoid resyllabification and refer only to output. However, they generally assume that aspiration is tied to coda position and add constraints that compare one output form to another. Kenstowicz (1996) extends correspondence constraints from Input-Output and Base-Reduplicant relationships (McCarthy and Prince 1995) to Uniform Exponence. This constraint requires lexical items that do not occur independently to match their own output shape in other forms, as in (14).

(14) UNIFORM EXPONENCE (UE): minimize the differences in the realization of a lexical item

Kenstowicz uses UE to compare the realization of the prefix /des/ before a vowel-initial word with its realization before a consonant-initial one. Since /des/ does not occur independently, the interpretation of the constraint is problematic; how does the prefix know to match its realization before a consonant-initial word, rather than vice versa? Furthermore, to avoid applying UE to pairs with suffixes ([meh], *[meheh] for 'month, months'), Kenstowicz claims that UE is defined for specific morphemes only, rather than for all prefixes or words. Colina (1997) attempts to improve this account by using a prosodic and a segmental version of UE and claims UE need not be specified for prefixes. For dialect differences, however, she notes that UE might be reranked or use a different allomorph as a basis for comparison in different dialects, again revealing that there is no principled basis for the choice of allomorph on which the constraint is evaluated.

Face (1999) avoids stipulating that UE as applies only to specific morphemes by analyzing prefixes as independent phonological domains (PDs), based on the observation that prefixes act differently from inflectional suffixes in syllabification. He modifies UE to the constraint UE-PD (15), and claims that UE-PD forces prefix-final /s/ to match the form of the prefix in preconsonantal position, where aspiration is forced by some high ranking constraint.

(15) UE-PD: minimize the differences in the realization of a phonological domain (Face 1999)

However, the PD analysis gives prefixes the same status as words, leaving their lack of stress to stipulation. Apart from problems with the term PD, the problem of dialect variation remains. Face's analysis requires that in dialects such as Caribbean I, prefixes are analyzed as PDs so that prefix-final /s/ will aspirate. Equating prefixes and words predicts that aspiration word-finally is accompanied by aspiration prefix-finally. This is not the case in dialects such as Rio Negro and Caribbean II. Also, as Face explains both syllabification and aspiration by 'PDs', changing their structure for different aspiration facts would also result in different syllabification facts. Such an account is therefore unable to simultaneously handle the similarities as well as the differences in different dialects.

In my analysis, I have used independently motivated constraints on prosodic structures to account for both syllabification and dialect variation. There are several advantages to this account. I have captured the different locations of aspiration and velarization in different dialects with a uniform syllabification and uniform prosodic word structure for all varieties of Spanish. We do not need rule ordering and abstract syllabifications, which lead to opacity and which require derivational levels rather than a fully parallel analysis. Finally, the interpretation of prosodic constituent integrity as alignment allows for an OT account without questionable extensions to UE constraints.

---

[1] I thank the audience at LACUS 2000 and anonymous reviewers for useful comments and Irene Moyna for additional examples from Spanish. I take full responsibility for any remaining errors of data or analysis.

[2] In the data that follows, ' . ' is used for syllable boundaries, '#' for word boundaries, and '+' for morpheme boundaries. Accent marks have been omitted throughout the paper, as the analysis here does not address stress. In (2)–(5), clusters or syllables of particular interest are underlined. For onset phonotactics, see Colina (1997).

[3] As a reviewer rightly points out, the use of the terms left and right are artificial when applied to spoken languages; the edges being referred to are the beginning and end of the unit (syllable, word, utterance). As 'left' and 'right' are entrenched in the theory, I use them here for compatibility with other work in OT.

[4] In the OT tableaux, '*' is used to indicate a violation of the constraint above the column in which it appears. '!' indicates that a candidate output has lost because of that constraint violation. A violation of a constraint in a higher ranked column is worse than one in a lower ranked column, and multiple violations of the same constraint are worse than single violations; either can cause a candidate to lose to a more optimal candidate.

[5] Nuñez-Cedeño and Morales-Front (1999) independently came to a similar description of these facts, though they only sketch the OT principles involved and do not develop the difference between prefixes and suffixes.

[6] The reader is referred to her work for details. Peperkamp claims only to show that monosyllabic prefixes do not incorporate into the PW of the base, but not whether they are PWs themselves in Spanish. I argue that the difference among dialects in their treatment of prefix-final and word-final /s/ supports a recursive structure as correct for Spanish.

[7] I diverge from Peperkamp's analysis by allowing word and syllable boundaries to be misaligned, a violation of PROPERNESTING (Peperkamp 1997). See Wiltshire (1999) for arguments that this constraint should be violable.

[8] The constraint is a combination of two constraints, one aligning left edges of PWs and lexical words, and one aligning the right edges. The two need not be ranked independently here, so they have been combined, though I will indicate by the asterisk which edge is responsible for the violation.

[9] The prefixes ending in *-s* are *bis-*, *des-*, *dis-*, *es-*, *ex-*, *minus-*, *plus-*, *pos-* and *tra(n)s-*. I use *des-* here as representative.

[10] Although early accounts of /s/ aspiration use preconsonantal position (Guitart 1979), recent accounts use coda position, following Harris (1983). Dialects such as BAA (Kaisse 1998), in which pre-pausal /s/ is [s], result naturally from a preconsonantal description, but require an additional constraint if aspiration is coda-based. Though special constraints for phrase-final position are possible, this option is not pursued here, pending further research.

REFERENCES

COLINA, SONIA. 1997. Identity constraints and Spanish resyllabification. *Lingua* 103:1–23.

FACE, TIMOTHY. 1999. Reexamining Spanish 'resyllabification'. Paper presented at the 29th Linguistic Symposium on Romance Languages, Ann Arbor, Michigan, April 1999.

GUITART, JORGE. 1979. On the true environment for weakening and deletion in consonant-weak Spanish dialects. Paper presented at the Conference on Non-English Language Variation in the Western Hemisphere, University of Louisville.

HALE, MARK, MADELYN KISSOCK & CHARLES REISS. 1998. Output-output correspondence in Optimality Theory. WCCFL XVI.

HARRIS, JAMES W. 1983. *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge MA: MIT Press.

———. 1993. Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. In *The view from Building 20*, ed. by Ken Hale & S. Jay Keyser, 177–94. Cambridge: MIT Press.

HUALDE, JOSÉ IGNACIO. 1992. On Spanish syllabification. In *Current studies in Spanish linguistics*, ed. by Hector Campos & Fernando Martínez-Gil, 475–93. Washington DC: Georgetown.

KAISSE, ELLEN. 1998. The prosodic environment of s-weakening in Argentinian Spanish. In *Grammatical theory and Romance languages*, ed. by Karen Zagona, 123–34. Amsterdam: John Benjamins.

———. 1999. Resyllabification precedes all segmental rules: evidence from Argentinian Spanish. In *Formal perspectives on Romance linguistics*, ed. by M. Authier, B. Bullock & L. Reed, 197-210. Amsterdam: John Benjamins.

KENSTOWICZ, MICHAEL. 1996. Base-identity and Uniform Exponence: Alternatives to cyclicity. In *Current trends in phonology: Models and methods, vol 1.*, ed. by J. Durand & B. Laks, 363–93. Salford: ESRI.

MCCARTHY, JOHN & ALAN PRINCE. 1995. Faithfulness and reduplicative identity. In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk, 249–384. Amherst: GLSA.

NÚÑEZ CEDEÑO, RAFAEL A. & ALFONSON MORALES-FRONT. 1999. *Fonología generativa contemporánea de la lengua español.* Washington DC: Georgetown University Press.

PEPERKAMP, SHARON. 1997. *Prosodic words*. University of Amsterdam PhD dissertation. [Den Haag: Holland Academic Graphics, HIL dissertations: 34.]

SELKIRK, ELISABETH. 1984. On the major class features and syllable theory. In *Language sound structures*, ed. by M. Aronoff & R. Oehrle, 107–36. Cambridge MA: MIT Press.

———. 1995. The prosodic structure of function words. In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk, 439–69. GLSA: Amherst.

WILTSHIRE, CAROLINE. 1999, Review of *Prosodic words*, by Sharon Peperkamp. *Glot international* 4.4:16–18.

———. in press. Variation in Spanish aspiration and prosodic boundary constraints. In *Selected papers from the XXIXth Linguistic Symposium on Romance Languages*, ed. by D. Cresti, T. Satterfield & C. Tortora. Amsterdam: John Benjamins.

◆